

University of Sheffield

Visual Localisation for Pipe Inspection Robots using Prior Information of the Environment



Sarah Edwards

Supervisors: Sean Anderson, Jonathan Aitken

A report submitted in partial fulfilment of the requirements
for the degree of Master of Philosophy

in the

Department of Automatic Control and Systems Engineering

December, 2023

Abstract

The ability of autonomous vehicles to explore and navigate their environments has improved greatly in recent years. However, many challenges remain to be overcome and many environments exist that present unique challenges to the problem of autonomous navigation. One such environment is subterranean pipe networks, such as sewers. There is a necessity to inspect these environments for maintenance purposes and, due to the difficult and time consuming nature of manual inspections, a desire to automate the process with mobile robots. As such, the challenges presented by subterranean pipe networks to autonomous navigation must be overcome.

This thesis contains a review of the existing research on SLAM algorithms with a focus on their use in pipe networks. The review aims to understand the state of the art approaches to solving the SLAM problem and how those approaches may be applied to pipe networks. The included review also highlights the specific properties of pipe networks that present additional challenges to SLAM algorithms that may not be encountered in more general environments. Of those presented in the review section, the principle challenges this thesis is concerned with stem from the relatively featureless and uniform pipe walls and the lack of communication with external systems. These characteristics make the feature detection and matching used by most visual SLAM algorithms difficult and the positional correction provided by systems such as GPS impossible.

This work attempts to overcome these challenges by exploiting the predictability of landmarks within pipe environments and the ability to know or create maps of those landmarks prior to navigation within the pipes themselves. This is presented in two publications, the first of which details a method of visually recognising joints and manholes and associating them with their known locations in a prior map to perform odometry. This approach has led to a navigational system able to achieve a Mean Absolute Error up to 6 times lower than a state of the art algorithm in the same environment and which is capable of operating in environments

where the state of the art algorithm fails entirely.

The second publication presents a system of further robustification in full networks by representing the results of an odometry system as a particle filter loosely constrained to a simple map. The system accounts for uncertainty in the path taken by the agent within the map by instantiating multiple particle filters representing different hypotheses and then assessing their probability of being the true trajectory based on the odometry outputs and the map. This allows the system to recover from failure and also improves the mean position and heading errors of the original odometry system by an order of magnitude and the final positional error by three orders of magnitude.

These publications represent a step toward the autonomous inspection of pipe networks and demonstrate the effectiveness of taking non standard and more specialised approaches over more commonly used and generalised methods when attempting autonomous navigation in uniquely challenging environments.

Contents

1	Introduction	1
1.1	Background and Motivation	1
1.2	Aims and Objectives	2
1.2.1	Thesis structure	2
1.3	Contributions arising from the thesis	3
1.3.1	First author publications	3
1.3.2	Publication contributions	3
2	Review of SLAM for Pipe Inspection Robots	4
2.1	Summary of Research	4
2.2	Publication	6
3	Visual Odometry Using Joint and Manholes Detections	33
3.1	Summary of research	33
3.2	Publication	35
4	Multiple Hypothesis Particle Filtering For Visual Odometry	49
4.1	Summary of research	49
4.2	Publication	50
5	Conclusions	56

Chapter 1

Introduction

1.1 Background and Motivation

Water infrastructure receives a significant amount of investment and research (Government Office for Science, 2012) and the inspection of buried pipe networks is an important part of maintaining that infrastructure (Rizzo, 2010). A shadowing of water company employees revealed that, currently, pipes are inspected using tethered vehicles (WRC Infrastructure, 2020) mounted with cameras and the footage is manually searched for damage and blockages. This is a time consuming process, requiring multiple workers to visit the entirety of a network in order to perform a complete inspection (Wirahadikusumah et al., 1998). Advances in mobile robotics, detailed in chapter 2, have presented an opportunity to replace these manual inspections with continuous observation of pipe networks by autonomous robots. This opportunity, however, is not a trivial one to take. Pipe networks present many challenges that make autonomously navigating in them difficult even for algorithms that show impressive results in other environments, such as those developed for road vehicles (Aitken et al., 2021).

Two of the major challenges to the autonomous navigation of pipe networks are a lack of environmental features, with those features that do exist exhibiting low visual diversity, and no or limited communication with the surface (Aitken et al., 2021). However, pipe networks also possess some exploitable aspects, such as a somewhat predictable environment with stringent constraints on where a vehicle within the network can move (Twort et al., 2009 and Butler et al., 2019). Despite only manholes being visible on the surface, the predictability of pipe networks makes mapping them from above ground feasible (Yatim et al., 2014). It was found

during the course of this research that the connections between manholes can be modelled as one dimensional paths and the networks are designed to allow pipe contents to flow in one direction towards a single point, a collection of identified manholes can be represented as a directed tree graph with its edges representing pipes, this is similar to map representations for pipe localisation used in (Worley et al. 2019 and Worley et al. 2020). The exploitation of a pipe networks ability to be represented in this way, along with the predictability of its landmarks, forms the basis for the work undertaken for this thesis.

1.2 Aims and Objectives

The research presented in the following chapters aims to develop autonomous visual odometry methods for pipe network navigation by leveraging prior knowledge of the environment and exploiting the environmental features that cause conventional methods to under perform or fail.

To do this the following objectives must be met:

- Creating a representation of the environment from existing or gathered data that can be used as a map by an autonomous agent (chapters 3, 4 and 5).
- Developing a method of identifying predictable landmarks from camera footage (chapter 3).
- Automatically reconciling detected landmarks with those held in the prior map in order to locate the agent relative to those landmarks (chapter 3).
- Determining a path between the identified landmarks that accounts for the constraints placed on movement by the environment whilst also allowing for errors in both the prior map and the landmark detections (chapter 3 and 4).
- Robustifying the process so that an agent can reliably navigate a network of interconnected pipes where there is uncertainty surrounding the path taken when multiple are available (chapter 4).

1.2.1 Thesis structure

Chapter 2 presents a literature review of the wider aspects of robot localisation and mapping for sewer and water pipes (Aitken et al., 2021).

Chapter 3 presents the research conducted on the detection of pipe joints and manholes from below ground using camera data and the use of those detections for odometry. This chapter includes a summary of research conducted followed by a journal publication (Edwards et al., 2023).

Chapter 4 presents the reconciliation of odometry data with a map prior using multiple hypothesis particle filtering. This chapter includes a summary of research followed by a conference publication (Edwards et al., 2021).

Chapter 5 contains a reflection on the work included in chapters 2 and 3 as well as a brief outline of possible continuations of this research.

1.3 Contributions arising from the thesis

1.3.1 First author publications

1. Edwards S, Zhang R, Worley R, Mihaylova L, Aitken J and Anderson SR (2023) A robust method for approximate visual robot localization in feature-sparse sewer pipes. *Front. Robot. AI* 10:1150508. doi: 10.3389/frobt.2023.1150508
2. Edwards, S., Mihaylova, L., Aitken, J.M., Anderson, S. (2021). Toward Robust Visual Odometry Using Prior 2D Map Information and Multiple Hypothesis Particle Filtering. In: Fox, C., Gao, J., Ghalamzan Esfahani, A., Saaj, M., Hanheide, M., Parsons, S. (eds) *Towards Autonomous Robotic Systems. TAROS 2021. Lecture Notes in Computer Science*, vol 13054. Springer, Cham. https://doi.org/10.1007/978-3-030-89177-0_19

1.3.2 Publication contributions

1. J. M. Aitken et al., "Simultaneous Localization and Mapping for Inspection Robots in Water and Sewer Pipe Networks: A Review," in *IEEE Access*, vol. 9, pp. 140173-140198, 2021, doi: 10.1109/ACCESS.2021.3115981.
2. R. Zhang, R. Worley, S. Edwards, J. Aitken, S. R. Anderson and L. Mihaylova, "Visual Simultaneous Localization and Mapping for Sewer Pipe Networks Leveraging Cylindrical Regularity," in *IEEE Robotics and Automation Letters*, vol. 8, no. 6, pp. 3406-3413, June 2023, doi: 10.1109/LRA.2023.3268013.

Chapter 2

Review of SLAM for Pipe Inspection Robots

2.1 Summary of Research

During the initial stages of this project, a literature survey was conducted to review existing research on the problem of autonomous pipe inspection as well as exploring relevant research in surrounding fields and determining the current state of the art. This information was collated into the publication below which constitutes a review of the literature relevant to this thesis.

The publication begins by motivating the need for autonomous robotic inspection of pipe networks and giving a broad overview of the existing research on the SLAM (simultaneous localisation and mapping) problem in pipe networks. It then details the challenges faced when attempting to perform SLAM in pipe networks and explores the general requirements of the water industry and translates those requirements into technical requirements any proposed SLAM system must meet in order to operate in pipe networks. The review asserts that future inspection of buried pipe networks will be carried out by small autonomous robots. It finds that, in addition to the general difficulties of performing SLAM, buried pipe networks suffer from a lack of external communication and erroneous prior information, as well as being dark, narrow, difficult to traverse, and lacking much of the visual and environmental information used in more generalised SLAM algorithms.

The publication then provides multiple sections constituting a detailed review of existing research related to the problem. This begins with an overview of SLAM methods and is followed

by a review of map representation methods and their applicability to pipe networks then an exploration of how different sensors can be used for SLAM in pipes.

In brief, these sections find that a variety of approaches to SLAM exist and are well explored, however they are generally too computationally expensive to be able to be performed on robots meeting the size constraints highlighted earlier in the review. It posits that robots in pipes will likely perform only localisation using existing maps and will perform full SLAM only occasionally. It is also, however, found that existing pipe network maps are frequently inaccurate or incomplete and are intended to be read by humans. A number of ways of representing maps are discussed in the review, and maps that contain a spatial representation of explicitly classified features (semantic maps) are highlighted as potentially suitable for pipe networks due to them containing many already classified landmarks. The map section concludes by suggesting a hybrid map representation that utilises differing scales and levels of abstraction based on the task being performed may be ideal for use by autonomous robots in pipes.

Finally, the review of sensor methods outlines that the power and size restrictions imposed by the operating environment will limit the type and number of sensors able to be used by robots in pipe networks. The review suggests that cameras are an effective method of localising via the detection of landmarks but highlights their dependence on light, the source of which must be carried on the robot in pipe networks. Laser scanners and Lidar do not require light, however their increased size and expense over cameras as well as the lack of features used by range based pose estimation algorithms is stated as a concern. Additionally, acoustic sensing is highlighted as being well suited for pipe environments, however its need for a fixed source of sound would limit the mobility of the robot. Similarly to the findings on map representations, the review posits that the use of multiple types of sensors will likely be the ideal approach to performing SLAM in pipe networks.

The review concludes by highlighting the unsolved challenges surrounding the problem of autonomous inspection of pipe networks, outlining the need for any solutions to be robust and trustworthy, highlighting the lack of research into the use of prior map information in SLAM, and detailing the potential usage of active SLAM and multi-robot SLAM in pipe networks, both of which were found to not yet have been researched for this purpose.

The author of this thesis contributed material throughout the publication and particularly relating to visual SLAM loop closing (section IV-B 2), visual odometry (section VI-B) and the

use of prior maps (section VII-B).

2.2 Publication

Received 6 August 2021, accepted 8 September 2021, date of publication 27 September 2021, date of current version 19 October 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3115981

Simultaneous Localization and Mapping for Inspection Robots in Water and Sewer Pipe Networks: A Review

JONATHAN M. AITKEN¹, MATHEW H. EVANS¹, ROB WORLEY¹, SARAH EDWARDS¹,
RUI ZHANG¹, TONY DODD², LYUDMILA MIHAYLOVA¹, (Senior Member, IEEE),
AND SEAN R. ANDERSON¹

¹Department of Automatic Control and Systems Engineering, The University of Sheffield, Sheffield S1 3JD, U.K.

²School of Digital, Technology and Arts, Staffordshire University, Stoke-on-Trent ST4 2DE, U.K.

Corresponding author: Sean R. Anderson (s.anderson@sheffield.ac.uk)

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC), U.K., support through the Pervasive Sensing for Buried Pipes (Pipebots) Project under Grant EP/S016813/1.

ABSTRACT At the present time, water and sewer pipe networks are predominantly inspected manually. In the near future, smart cities will perform intelligent autonomous monitoring of buried pipe networks, using teams of small robots. These robots, equipped with all necessary computational facilities and sensors (optical, acoustic, inertial, thermal, pressure and others) will be able to inspect pipes whilst navigating, self-localising and communicating information about the pipe condition and faults such as leaks or blockages to human operators for monitoring and decision support. The predominantly manual inspection of pipe networks will be replaced with teams of autonomous inspection robots that can operate for long periods of time over a large spatial scale. Reliable autonomous navigation and reporting of faults at this scale requires effective localization and mapping, which is the estimation of the robot's position and its surrounding environment. This survey presents an overview of state-of-the-art works on robot simultaneous localization and mapping (SLAM) with a focus on water and sewer pipe networks. It considers various aspects of the SLAM problem in pipes, from the motivation, to the water industry requirements, modern SLAM methods, map-types and sensors suited to pipes. Future challenges such as robustness for long term robot operation in pipes are discussed, including how making use of prior knowledge, e.g. geographic information systems (GIS) can be used to build map estimates, and improve multi-robot SLAM in the pipe environment.

INDEX TERMS Water, sewer, network, pipe networks, robots, SLAM, data fusion, Bayesian estimation, visual odometry, laser and lidar scanning.

I. INTRODUCTION

Water is one of our most precious natural resources. Pipe networks transport water between sources and destinations, and similarly, sewer and drainage pipes transport waste products away from the customer to processing plants. Inspection and maintenance of water pipe networks [1] is crucial for maintaining a robust water supply and conserving the resource, and in the case of wastewater preventing contamination from leaking sewer pipes and removing blockages. In the UK, the buried pipe network for water and wastewater is around 0.8 million kilometres in combined length [2], whilst the

USA has 1.2 million miles of water supply mains and a similar amount of sewer pipes [3]. Investment in water and waste infrastructure is correspondingly large: over £250 billion is invested in UK water infrastructure [4], whilst the USA Environmental Protection Agency estimates that \$271 billion must be invested over the next 20 years for wastewater/storm-water upgrades and \$384 billion for drinking water upgrades [5]. Failure in pipe networks, in terms of a pipe leak, burst or blockage, can lead to severe disruption, including loss of water supply and road closures whilst the damage is repaired. It is estimated that over 3000 million litres of water is lost to leaks every day in the UK [6], and about 900 billion gallons of untreated sewage is discharged into USA waterways each year [7]. Therefore, continuous inspection, monitoring of

The associate editor coordinating the review of this manuscript and approving it for publication was Saeid Nahavandi¹.

deterioration, and detection and localization of damage in water and sewer pipes is of utmost importance.

Currently, there is no mobile robot technology in real world use in industry that can autonomously monitor pipes to detect and localise defects over a large spatial scale and a long time duration. Industry methods used to monitor the pipes tend to be manually operated, tethered systems, such as in-pipe closed-circuit television (CCTV) inspection for sewer pipes [8] and above-ground technologies including ground penetrating radar [9], [10] and electromagnetic location (EML) [11] for both water and sewer pipes (and other underground infrastructure) [12].

Robotic devices have been developed in industry for water pipe inspection. Sahara [13], which is a tethered device and Smartball [14] are some of the most popular ones. Smartball is a newer, untethered technology that relies on a free-flowing (rolling) inspection. However, these two robotic technologies tend to be limited to specific short sections of pipe for one-off inspections.

Previous surveys focus on inspection in water pipes [15], [16], sewer pipes [17], small diameter pipes [18], and briefly on mapping for underground pipes [19], but the main emphasis is on the types of sensors and surveying technologies depending on the types of pipelines. There is a gap for a focused survey on state-of-the-art robot localization and mapping in water and sewer pipe networks, which is addressed here.

Robot mapping and localization methods are usually based on simultaneous localization and mapping (SLAM), where raw sensor data is processed in a front-end for feature extraction, data association and loop closing, and robot location and map estimates are produced by a back-end optimization algorithm. SLAM is useful for both robot control and navigation, as well as in the context discussed here of mapping the pipe network and localizing defects. The SLAM problem and its classical solution methods are well reviewed in [20], [21] and more recent approaches and future challenges are presented in [22]. The aim of this paper is to survey the state of the art in SLAM in pipe networks. The focus is on water and sewer pipes but we also refer to other pipes (e.g. gas) where relevant work has been done.

The remainder of the paper is structured as follows. In section II we describe the particular challenges of the pipe environment for SLAM. In section III we describe the high-level requirements of the water industry and relate this to technical requirements for SLAM in pipes. In section IV we provide an overview of methods for SLAM. Section V reviews map representations and discusses their suitability for pipes. Section VI reviews different sensors used for SLAM in pipes. Future challenges are considered in section VII. Finally, conclusions are given in section VIII.

II. CHALLENGES FOR SLAM IN WATER AND SEWER PIPES NETWORKS

SLAM is fundamentally a difficult problem to solve because to estimate a map you need a good estimate of your location,

and to estimate your location you need a good estimate of the map. Odometry methods can be used to estimate location but the estimate drifts over time. SLAM is particularly challenging therefore when exploring new areas, or areas that lack discriminative features used to recognise places to correct odometry drift. To function reliably, a SLAM system should be robust in all of its components - front-end sensing (where it is generally thought that the use of multiple sensor-types improves robustness), landmark recognition, and back-end optimization algorithms that produce the location and map estimates. These are general challenges for SLAM but there are a number of specific challenges for SLAM in water and sewer pipes that are described in this section (and summarised in Table 1).

Water and sewer pipes are typically buried underground, meaning that robots in these pipes cannot receive GPS signals to estimate their location. GPS is one of the most popular and standard methods of localization, and is also commonly used for drift correction of dead reckoning sensors such as those based on odometry and inertial sensing. The lack of GPS signals makes the problem of SLAM in pipes far more challenging than typical outdoor scenarios.

Water and sewer pipes are often relatively small in diameter, with the majority of water pipes worldwide in the region of 100-150 mm in diameter, and the majority of sewer pipes in the UK of 300 mm or less in diameter (see section III for more details). This limits the size of a robot, particularly its sensor payload, computational hardware and batteries. This means that a typical pipe robot is likely to be limited to a small number of sensors and that certain types of sensor might be unsuited to the environment, for instance those that are not readily miniaturised or consume large amounts of power.

The inside of water and sewer pipes are difficult environments to navigate through: they are dark, water-filled (always in water distribution pipes and with time-varying levels of waste-water in sewer pipes), with possible occlusions of sensors occurring in sewers from waste products. This particularly complicates SLAM based on vision, because a light source is needed. The possibility of the dirty environment fouling the sensors also raises the possibility of sensor failure, so robustness in sensing and navigation, and failure-aware robotics is a critical issue.

Mobility also presents a challenge because water pipes are pressurised and it may be difficult for a robot to move against the flow, whilst sewer pipes move flow down the gradient and consequently connections in manholes can often have significant drops in height at the inlet pipe, from the centimetre scale up to metres. So, for both water and sewer pipes it is easier to move in one direction (with flow, or down the gradient respectively), which can impact active SLAM methods, where the goal is to actively explore and map the environment.

An additional challenge is the lack of accurate maps of pipe networks. Whilst water utilities do usually possess pipe network maps, in the form of top down, 2D line drawings (see Fig. 1), they can often contain errors, due to:

TABLE 1. Specific challenges for SLAM in water and sewer pipes beyond general challenges for SLAM.

Problem	Description
1. GPS-denied	GPS signals do not transmit underground to buried water and sewer pipes so GPS cannot be used for absolute localization and to correct drift in odometry methods.
2. Small diameter pipes	Water and sewer pipes tend to be of small diameter (often 150 mm for water, 300 mm or less for sewers), limiting the size of the robot, including sensors, computational hardware and batteries.
3. Dark, wet/dirty environment	Pipes are dark environments, so a light source is needed for vision methods; water pipes are always full of pressurized water, whilst sewer pipes have time-varying amounts of (dirty) fluid that can interfere with sensors.
4. Flow and gradient	Water pipes have significant flow rates, whilst sewer pipes move waste down the gradient so connections can have severe drops, from centimetres to meters - both of these characteristics mean that travel is usually much easier in one direction (with flow/down the gradient), which can be a problem for mobility and affect exploration e.g. in active SLAM.
5. Prior map errors	Water and sewer pipe network maps tend to exist and be available but also tend to contain errors, mainly due to discrepancies arising between <i>planned</i> installation and <i>actual</i> installation, which poses a problem for over/under reliance on prior maps used for localization.
6. Restricted movement and sensing	Pipes exist in a 3D environment but the robot is restricted in movement to effectively 1D along the enclosed environment of the pipe - this means that usually only nearby landmarks can be observed and only from a limited perspective, which may affect accuracy and convergence in SLAM algorithms.
7. Feature sparse environment	The inside of a pipe often lacks distinctive features that are used in SLAM for tasks such as pose estimation. In pose estimation, features are matched over successive frames (with cameras) or scans (with laser scanners). Lack of features can cause these algorithms to fail.
8. Perceptual aliasing	Different locations in the pipe network can generate a similar perceptual footprint because the visual appearance and geometry can be very similar at completely different locations, which presents a problem for loop closing.

**FIGURE 1. An example of a typical water distribution map for a UK town. The scale is approximately 2 km from top to bottom.**

- 1) discrepancies arising between planned replacement (upon which maps are usually based) and actual installation on-site,
- 2) loss of records (pipes can be tens of years old),
- 3) lack of precise record keeping.

Errors in a map can be misleading to a SLAM system, potentially causing incorrect data associations, that can in turn affect a fragile back-end estimation algorithm causing the whole system to fail. Therefore, a key challenge is to how to incorporate prior map knowledge into the system in ways that will be both beneficial and robust.

SLAM algorithms usually rely on reducing uncertainty in the map and robot location by taking successive measurements of the same (static) map features from different positions. However, the robot movement in pipes is restricted to predominantly one dimensional (1D) movements along the pipe, whilst the pipe network itself exists in a three

dimensional (3D) space. This restricted movement means that the same map features cannot be observed from many perspectives to help reduce map and localization uncertainty.

The inside of a pipe tends to be a feature sparse environment, lacking distinctive landmarks that can be reliably and repeatedly detected. This is a particular problem for sensors that observe the external world, like cameras and laser scanners, although sensors based on internal measurements of motion such as inertial measurements and wheel odometry are unaffected by this problem. For instance cameras are often used for feature-based visual odometry, where features common to successive image frames are used to estimate the robot pose. Insufficient features can cause such an algorithm to fail.

Loop closing is an important component of a SLAM system, which refers to recognising places when the robot returns to a previously visited location. Loop closure enables drift errors to be corrected and so improves the accuracy of a SLAM system. Loop closing is likely to be more challenging in pipes because it is a highly homogeneous environment with little variation in visual features, and similar types of structure and geometry repeated throughout the network, i.e. the cylindrical shape of pipes and the standard shapes of e.g. pipe joints, junctions and manholes. The pipe environment is therefore prone to perceptual aliasing, where different places in the environment generate a similar perceptual footprint. This is a challenge for SLAM in pipes because the robot might be prone to false positives in loop closure, where the robot mistakenly recognises a place and closes a loop, and false negatives, where the robot fails to recognise that it has returned to a previously visited location, preventing successful loop closures.

Finally, it is worth noting that amongst the various challenges for SLAM in pipes, there are also some advantages

TABLE 2. Accuracy levels designated in the British Standards Institution PAS128: Specification for underground utility detection, verification and location.

Accuracy Level	Positional Accuracy	Applies to	Comment
1	± 25 mm Vertical	Z data	Coincides with PAS128 quality level A
	± 50 mm Horizontal	X, Y data	
2	± 50 mm	X, Y, Z data	As Level 1, but relaxes accuracy in vertical direction
3	± 150 mm	X, Y, Z data	Achievable using GPS and RTK methods
4	± 300 mm	X, Y, Z data	Achievable by post-processed mapping grade GPS
5	± 1000 mm	X, Y, Z data	Achievable with mapping grade GPS.
6	± 1000 mm	X, Y data	Accuracies of Z data are unreliable/not available.
7	Indeterminate		Accuracies of X, Y, Z data are indeterminate.

TABLE 3. Utility detection quality levels as defined by the American Society of Civil Engineers.

Quality Level	Comment
A	Found using surface geophysical methods, to an accuracy of 15 mm
B	Found using surface geophysical methods
C	Found using assets observable above ground
D	Found using previous surveys

as well. For instance, it is typically the case that sewer pipes are laid in straight lines and that changes of direction occur at points such as manholes, which should help avoid drift in heading estimates. Also, manholes for sewers, and fire hydrants for water pipes, occur relatively frequently (approximately tens of metres apart) and can be accurately mapped from above-ground, which then provides known reference points when correctly recognised and data associated from within the pipe. Therefore, there are some aspects of the environment that can be exploited to simplify the SLAM problem.

III. WATER INDUSTRY REQUIREMENTS FOR PIPE NETWORK MAPPING AND DEFECT LOCALIZATION

In this section we consider water industry requirements and technical requirements for robot mapping and localization in pipe networks. Specifications quoted in the following section of high-level and low-level requirements are guided by published literature (discussed in the next sections) and interactions with key stakeholders through personal correspondence and knowledge sharing events. These stakeholders include water utilities' representatives and water industry technology companies, which in turn interact, and are guided by, other stakeholders such as the customers, regulatory bodies and the government.

The regulatory body for the UK, the Water Services Regulation Authority or OFWAT, has published requirements that lists resilience, and in particular, *operational resilience*, as a key requirement [23]. Operational resilience means reducing the probability of water supply interruptions and wastewater flooding, as well as mitigating the impact of any disruption through efficient handling, good communication and quick recovery.

The requirement of operational resilience in the water industry has many different contributing factors but the focus of this survey is primarily on two key aspects that robot mapping and localization can aid with: 1. pipe network mapping, so that the water utilities know where their assets are located and 2. defect localization, so that water utilities know where

to target repairs, especially when this incurs the cost and disruption of excavating in the street.

A. PIPE NETWORK MAPPING

The locations of buried pipes are usually not fully known by the water and sewer companies responsible for managing pipe networks. This can be due to a number of reasons, such as the pipe locations not being recorded during installation, or the information not being recorded accurately, or the information being lost over time. Therefore pipe network mapping is an essential task for robot inspection systems so that water utilities know where their assets are located.

A pipe network map can be described using a few key variables:

- 1) pipes coordinates in 3 dimensions, i.e. X , Y , Z positions,
- 2) pipe diameter d ,
- 3) pipe gradient, g .

Surveys for existing methods already have designated accuracy levels that in-pipe robots will have to compete with. For instance the British Standards Institution (BSI) has produced the publicly available specification (PAS) *PAS 128* [12], for underground utility detection, verification and location, which has the accuracy levels specified in Table 2. Similarly, the American Society of Civil Engineers (ASCE) has produced the *Standard Guideline for the Collection and Depiction of Existing Subsurface Utility Data* [24], which gives four utility detection quality levels, shown in Table 3.

B. PIPE DEFECT LOCALIZATION

There are two main type of defects that need to be localised in pipe networks:

- 1) minor defects - knowing their location is important to monitor the defect over time,
- 2) major defects - finding their location is essential for the repair process - either conventionally by excavation or using trenchless technologies.

TABLE 4. Water distribution pipe diameters for typical city/urban/rural environments from the UK, South East Asia and the USA (Data obtained from Table 13.1 in [25]).

Diameter		UK		South East Asia		USA	
		city/urban	urban/rural	city/urban	urban/rural	city/urban	small urban
mm	inches	%	%	%	%	%	%
50-80	2-3	4.5	26.5	0.5	8.0	1.0	
100-150	4-6	73.0	49.7	50.1	62.5	42.0	44.0
200-225	8-9	9.9	7.9	21.3	7.5	24.0	30.0
250-350	10-14	6.1	6.0	22.5		16.0	
375-450	15-18	1.9	3.8	0.4	14.0	4.0	26.0
500-600	20-24	0.8	3.3	3.0		4.0	
Over 600	Over 24	3.8	2.8	2.3	4.0	9.0	
		100.0	100.0	100.0	100.0	100.0	100.0

Excavation size should be kept to a minimum to reduce costs (including reinstatement), minimise disruption and limit the potential to damage adjacent buried utilities. Discussions with water utilities suggest a sub-0.5 metre accuracy is desirable for locating defects, to guide excavation, which defines an accuracy requirement for a robot SLAM system. One key point though, emphasised by industry, is that there is a preference for excavating in the correct location at the first attempt - i.e. a single larger excavation would generally be preferred to excavating multiple smaller holes at incorrect locations. This means that the SLAM system should correctly characterise its location uncertainty via probabilistic methods.

The time-scale of reporting defects and subsequent repair varies depending on the type of fault and the associated severity. Mapping the pipe network can be dealt with over a long time-scale (months or years) and will be continually ongoing. Blockages in sewer pipes and small leaks in water pipes can be dealt with also over a medium time-scale (weeks-months). Bursts in water pipes need to be dealt with in a short time-scale (days).

C. SIZE, WEIGHT, POWER, AND COST (SWaP-C) REQUIREMENTS

The size, weight, power and cost (SWaP-C) requirements of robotic systems for water and sewer pipe networks are an important consideration, and pertain to mapping and localization with regard to a number of issues. Size restrictions for robots, and sensors in particular, in small diameter pipes, limit the type of technology that can be used for SLAM. Power is restricted because the mobile robots will be untethered, and will have to travel between charging stations whilst mapping a network; it is essential they do not lose power between these points and become lost in the network, creating a blockage and additional cost of recovery. The cost of robot solutions for inspection and maintenance needs to be competitive with existing manual solutions, for example a 2 person crew with a manually operated CCTV inspection system.

The size of water and sewer pipes varies greatly. The majority of water distribution pipe diameters across the world tend to be in the range of 100-150 mm (see Table 4 for more details) [25], whilst very large trunk mains can be metres wide. Sewer pipes also vary greatly in size, with around 70% of sewers in the UK having a diameter less than 300 mm,

whilst only 9% have a diameter of 900mm or greater [26]. Therefore robots for either water distribution or sewer pipes will need to be relatively small to ensure coverage of a pipe network, or different sizes of robot will need developing. Small size robots will in turn necessitate small sensor payloads, batteries and computational hardware.

A crucial further detail is that the size of the entry point to the pipe network is a hard constraint on the robot size. In water distribution pipes fire hydrants provide natural access points, which would be attractive to use because they avoid costs associated with creating special access points for robots. New style through-bore hydrants in the UK have a pipe diameter of only 80mm, which gives an upper limit on robot size if these were to be used (although many older hydrants in the UK have sharp bends and restrictive valves which would make the insertion of a robot more difficult). Sewer pipes to a large extent avoid concerns over size for entry because they can be accessed by large manholes.

Power requirements relate mainly to actuators, sensors and computational hardware. Modern machine learning techniques that might be used in visual navigation algorithms, for instance based on deep learning [27], [28], might require relatively high power specialist computing devices (based on embedded general purpose graphics processing units GPGPUs). The power requirements, in terms of Watt-hours of operation, should also be considered in conjunction with the aim that these robots should ultimately perform long term inspection of the pipe network, over months and years. This means that the robots will require recharging, and hence there will need to be recharging points specially added to the pipe network. The question about power really becomes one of time and distance between charging, and the cost of infrastructure associated with installing charging points.

D. REQUIREMENTS SPECIFIC TO WATER DISTRIBUTION

Water is a unique commodity in pipes in that, unlike sewerage, oil, and gas (monitoring of pipes transporting the latter two has been reviewed elsewhere [29]), the water must be fit for human consumption. The use of robots must not degrade the safety of water, which creates two main issues in the robot design:

- 1) avoiding water contamination from introducing foreign bodies into the pipe network,

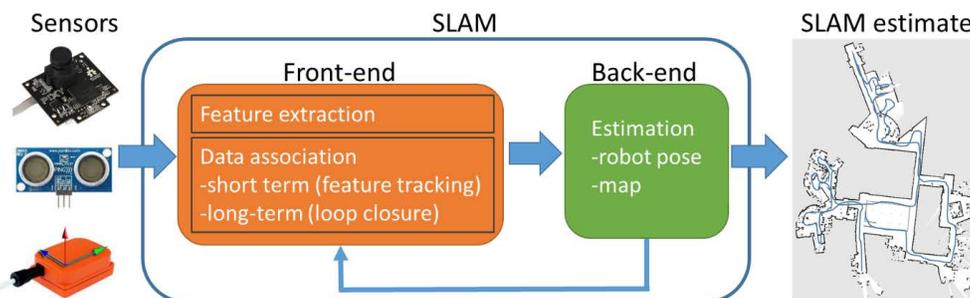


FIGURE 2. Typical SLAM system. The sensors transmit raw data to a ‘front-end’, which processes the raw data, extracts features and performs data association. The front-end transmits the processed data to the ‘back-end’, which estimates the robot pose (robot location and orientation) and the map. The back-end typically uses a standard method such as an extended Kalman filter, particle filter or smoothing method to perform the estimation. The back-end can provide feedback to the front-end for loop closure detection. (Redrawn and modified from [22]).

- 2) avoiding dislodging material from the pipe wall that will appear in the customer’s taps.

Whilst these requirements pertain more to mechanical design of the robot and the robot insertion device they are worth noting here due to their importance.

IV. OVERVIEW OF SLAM

The SLAM problem is usually thought of in two parts: the front-end and the back-end (see Fig. 2). The front-end processes raw sensor data to extract features and perform data association, i.e. feature tracking in the short term as a robot moves around an object and loop closure in the long-term when a robot returns to, and recognises, a location previously visited. The back-end part estimates the robot pose (its location and orientation in 2D or 3D space) and the map, using the information extracted by the front-end.

A. PRELIMINARIES

The typical SLAM problem is formulated by representing the history of robot poses $\mathcal{X}_{1:k}$, up to the current time-step k , together with the map \mathbf{m} , as the joint probability distribution

$$p(\mathcal{X}_{1:k}, \mathbf{m} | \mathcal{Z}_{1:k}, \mathcal{U}_{1:k}, \mathbf{x}_0), \quad (1)$$

where \mathbf{x}_0 is the initial robot pose, $\mathcal{U}_{1:k}$ is the history of control inputs (or odometry measurements) and $\mathcal{Z}_{1:k}$ is the set of all observed map landmarks, respectively

$$\mathcal{X}_{1:k} = \{\mathbf{x}_1, \dots, \mathbf{x}_k\}, \quad (2)$$

$$\mathcal{Z}_{1:k} = \{\mathbf{z}_1, \dots, \mathbf{z}_k\}, \quad (3)$$

$$\mathcal{U}_{1:k} = \{\mathbf{u}_1, \dots, \mathbf{u}_k\}. \quad (4)$$

The robot pose \mathbf{x}_k can be defined in 3D space as

$$\mathbf{x}_k = (X_k, Y_k, Z_k, \theta_k, \phi_k, \psi_k)^T, \quad (5)$$

where (X_k, Y_k, Z_k) defines the location of the robot in 3D space in the world coordinate frame and $(\theta_k, \phi_k, \psi_k)$ defines the orientation in terms of pitch, yaw, and roll, also in the world coordinate frame.

The map, \mathbf{m} , can be specified by the spatial locations of recognisable features or landmarks detected in the world, where

$$\mathbf{m} = (\mathbf{l}_1, \dots, \mathbf{l}_{n_m})^T \quad (6)$$

and $\mathbf{l}_i = (l_{i,x}, l_{i,y}, l_{i,z})^T$ is the location of the i^{th} landmark in the world coordinate frame, and n_m is the number of map features. The map can also be represented in a more dense form, e.g. as a grid map. We assume the map to be static throughout this paper, therefore \mathbf{m} is not a function of the time-step k , in contrast to the robot pose.

We assume that the robot state can be updated via

$$p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{u}_k) \iff \mathbf{x}_k = f(\mathbf{x}_{k-1}, \mathbf{u}_k) + \mathbf{w}_k, \quad (7)$$

where the $f(\cdot)$ function defines the state transition, \mathbf{u}_k is an input to the robot (or odometry measurement), and \mathbf{w}_k is state noise we assume to be Gaussian, zero-mean, white noise, with covariance \mathbf{Q}_k , i.e. $\mathbf{w}_k \sim \mathcal{N}(0, \mathbf{Q}_k)$.

Next, we assume that an observation \mathbf{z}_k of a map feature can be related to the state via the measurement function $h(\cdot)$,

$$p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{m}) \iff \mathbf{z}_k = h(\mathbf{x}_k, \mathbf{m}) + \mathbf{v}_k, \quad (8)$$

where \mathbf{v}_k is observation noise that we assume to be Gaussian, zero-mean, white noise, with covariance \mathbf{R}_k , i.e. $\mathbf{v}_k \sim \mathcal{N}(0, \mathbf{R}_k)$.

B. THE SLAM FRONT-END: FEATURE EXTRACTION, DATA ASSOCIATION, AND LOOP CLOSING

A variety of sensors are used in robotics, which produce different types of raw data. For instance cameras produce pixel data, whilst laser scanners produce range and bearing data. The SLAM front-end performs feature extraction initially, processing these differing types of raw data into a measurement format \mathbf{z}_k (the features) that can be used directly in the back-end estimation process. The front-end also performs data association, where a feature in the map \mathbf{m} is found which is most likely to be associated with the

measurement \mathbf{z}_k , which allows us to write the measurement equation above in (8).

Data association has two aspects:

- 1) short term data association, which handles the data association of consecutive sensor measurements;
- 2) long term data association, i.e. loop closing, where a robot recognises a place previously visited by associating current measurements to previously mapped landmarks.

Data association can have two main types of error:

- 1) false positives, where there is an incorrect association made between an observation and the map - this can lead to catastrophic failure in the back-end algorithm;
- 2) false negatives, where an observation is rejected as spurious - this leads to reduced data for the back-end, which can reduce the estimation accuracy but is arguably less serious than a false positive.

1) DATA ASSOCIATION METHODS

Data association can be performed using simple statistical validation gating, as used in target tracking [30]. The idea of a validation gate is that any previously mapped landmarks in the map have to fall within a region defined by the gate to be considered valid for association with the current measurement of a landmark. To clarify this point, consider a set of hypotheses $\mathcal{H} = \{j_1, \dots, j_{n_m}\}$ each of which associates one measurement \mathbf{z}_i with one landmark \mathbf{l}_{j_i} ; the measurement equation from (8) under hypothesis \mathcal{H} is therefore

$$\mathbf{z}_{\mathcal{H}} = h_{\mathcal{H}}(\mathbf{x}_k, \mathbf{m}_{\mathcal{H}}) + \mathbf{v}, \quad (9)$$

where $h_{\mathcal{H}} = (h_{1j_1}, \dots, h_{mj_m})$ is the collection of independent measurement models and $\mathbf{m}_{\mathcal{H}} = (\mathbf{l}_{1j_1}, \dots, \mathbf{l}_{mj_m})^T$ is the vector of map landmarks corresponding to $\mathbf{z}_{\mathcal{H}} = (\mathbf{z}_1, \dots, \mathbf{z}_{n_m})^T$. We can obtain a measure of distance between actual and predicted measurements under hypothesis \mathcal{H} by using the Mahalanobis distance,

$$D_{\mathcal{H}}^2 = (\mathbf{z}_{\mathcal{H}} - h_{\mathcal{H}}(\hat{\mathbf{x}}_k, \hat{\mathbf{m}}_{\mathcal{H}}))^T S_{\mathcal{H}}^{-1} (\mathbf{z}_{\mathcal{H}} - h_{\mathcal{H}}(\hat{\mathbf{x}}_k, \hat{\mathbf{m}}_{\mathcal{H}})) \quad (10)$$

where $\hat{\mathbf{x}}_k$ is the estimated robot pose, $S_{\mathcal{H}} = H_{\mathcal{H}} P_{\mathcal{H}} H_{\mathcal{H}}^T + R$ is the innovation covariance, $P_{\mathcal{H}}$ is the estimated covariance of robot pose and landmarks, and $H_{\mathcal{H}}$ is the Jacobian associated with the measurement function $h_{\mathcal{H}}$. For a measurement/landmark pairing to be considered acceptable (or jointly compatible), the Mahalanobis distance $D_{\mathcal{H}}^2$ should lie within the validation gate based on the chi-squared distribution,

$$D_{\mathcal{H}}^2 \leq \chi_{d,\alpha}^2 \quad (11)$$

where $d = \dim(h_{\mathcal{H}})$ is the dimension of the measurement function and α is the confidence level.

The gating method avoids unlikely data associations but a further problem is that multiple landmarks might fall into the gated region defined by (11), in which case there must be some additional method of data association. One of the simplest methods is to associate the measurement with the

nearest mapped landmark within the gate, i.e. a nearest neighbour approach, known as individual compatibility nearest neighbour (ICNN) [31].

The early implementations of SLAM used the nearest neighbour approach, e.g. [32], however, this method is prone to error when used with more than just a few landmarks [21]. Later approaches were developed that considered data association in a more robust batch mode, such as joint compatibility branch and bound (JCBB) [31]. JCBB has also been extended to handling 2D lidar scans where the measurements points are numerous and correlated [33].

Multiple hypothesis methods have also been used in data association for SLAM since very early work [34], to modern systems that e.g. extend JCBB to MHJCBB (multiple hypothesis JCBB) [35]. These methods have the potential to be more robust but also tend to be much more computationally intensive.

2) LOOP CLOSING BASED ON APPEARANCE RECOGNITION

Loop closing is often based on appearance recognition. This can be done using various sensors such as cameras and laser scanners. The fast appearance-based mapping (FAB-MAP) algorithm [36] developed a visual appearance recognition algorithm using a bag-of-words (BOW) approach modified from speech recognition - the idea is to build a database of images stored as numerical vectors in the BOW space. To construct the BOW, visual features are extracted from images of places to create visual words, using methods such as SIFT [37] or SURF [38], then each place is represented by a histogram, which is the frequency of occurrences of each visual word in the image. This approach was made more efficient in FAB-MAP2 enabling loop closures for much larger environments [39]. Appearance-based recognition using the BOW method has also been extended and made more robust by adding a fast geometrical check to the image matching procedure [40].

FAB-MAP only uses a single image to perform appearance recognition, which can be sensitive to variation in appearance, due to e.g. changing light conditions, seasonal changes, viewpoint variations and dynamic objects. Therefore, SeqSLAM was developed [41], which uses a sequence of images to perform appearance recognition and tends to be a more robust approach. SeqSLAM uses a sum-of-absolute differences (SAD) to match image sequences between the recent observations and a database. SeqSLAM has received various updates, such as more efficient versions that avoid exhaustive search and instead use efficient tree searching with nearest neighbours [42]. A review of these types of appearance-based recognition methods such as FAB-MAP and SeqSLAM can be found in [43].

Deep learning methods have now also been applied to the problem of visual appearance-based recognition, where initial approaches used deep convolutional neural networks (DCNNs) pre-trained for image recognition [44], [45]: the idea is to use the inner layers of the DCNN as automatically generated features rather than the ‘hand-crafted’

features typically used in appearance recognition such as SIFT or SURF. The development of NetVLAD improved the use of the basic DCNN by taking the DCNN features and passing them through a vector of locally aggregated descriptors (VLAD) module, specifically designed for image retrieval [46]. Subsequently, [47] introduced a large-scale dataset for purpose-specific training of DCNNs for visual place recognition, which demonstrated improvements on just using re-purposed image recognition DCNNs. Recent work has demonstrated that both traditional image processing features, histogram of oriented gradients (HOG), and DCNN features can give robust performance when using image sequences and maintaining multiple hypotheses for matching [48]. To combine advantages across the methods mentioned above, SAD, HOG and DCNN features have been fused to give state of the art performance [49].

Range data, often from laser scanners or lidar (light detection and ranging), can also be used to perform loop closing from appearance recognition either in 2D [50], [51] or 3D [52], [53]. The use of range data overcomes the sensitivity of cameras to different lighting conditions. Similarly to visual appearance recognition, deep learning has also been applied in recent years to the problem of range based 3D lidar appearance recognition [54].

C. THE SLAM BACK-END: ROBOT POSE AND MAP ESTIMATION

SLAM back-end estimation algorithms to obtain the robot pose, \mathbf{x}_k , and map, \mathbf{m} , fall broadly into one of two classes of algorithm:

- 1) Filter-based algorithms: these algorithms recursively estimate the current robot pose and map, i.e. they produce the estimate $\hat{\mathbf{x}}_k$, and are usually formulated as a Bayesian filtering problem. The main approaches are based on the extended Kalman filter (EKF), (i.e. EKF-SLAM, as used in early pioneering work [55]–[57] and then latterly with more rigorous convergence analysis [58]), the sparse extended information filter (SEIF) [59], the particle filter (PF) (e.g. DP-SLAM [60]), and the Rao-Blackwellised particle filter (RBPF), (e.g. FastSLAM [61], [62] and variants [63]).
- 2) Smoother/optimization-based algorithms: these algorithms estimate the history of robot poses and map using all the data in a batch mode, i.e. they produce the estimate $\hat{\mathcal{X}}_{1:k}$, and are usually formulated as a sparse nonlinear least squares problem. The main approaches are graph-based methods, such as GraphSLAM [64], [65], smoothing and mapping (SAM) [66] and incremental smoothing and mapping (iSAM/iSAM2) [67], [68].

To expand on the pose/map estimation problem, firstly, we define the following maximum *a posteriori* (MAP) problem following from (1),

$$\left(\hat{\mathcal{X}}_{1:k}, \hat{\mathbf{m}} \right) = \arg \max_{\mathcal{X}_{1:k}, \mathbf{m}} \left(p(\mathcal{X}_{1:k}, \mathbf{m} | \mathcal{Z}_{1:k}, \mathcal{U}_{1:k}, \mathbf{x}_0) \right), \quad (12)$$

where assuming that the measurements and state predictions are independent and from using Bayes rule we can say that,

$$p(\mathcal{X}_{1:k}, \mathbf{m} | \mathcal{Z}_{1:k}, \mathcal{U}_{1:k}, \mathbf{x}_0) = \eta \prod_{k=1}^N p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{m}) \prod_{k=1}^N p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{u}_k), \quad (13)$$

where η is a normalising constant. Note that to simplify the equations, we take the common assumption that the initial pose \mathbf{x}_0 is known with full certainty - even if this is not the case, in an arbitrary coordinate system, \mathbf{x}_0 can be taken to be at the origin and initialised to zero [69].

Substituting (7) and (8) into (13), which are both subject to Gaussian noise, and noting that maximising (13) is equivalent to minimising the negative log posterior, leads to the nonlinear weighted least squares problem, similar to that defined in [64] and [66],

$$J(\mathcal{X}_{1:k}, \mathbf{m}) = \sum_{k=1}^N \frac{1}{2} \|\mathbf{x}_k - f(\mathbf{x}_{k-1}, \mathbf{u}_k)\|_{\Lambda_k}^2 + \sum_{k=1}^N \frac{1}{2} \|\mathbf{z}_k - h(\mathbf{x}_k, \mathbf{m})\|_{\Sigma_k}^2, \quad (14)$$

where $\|\mathbf{a}\|_P^2 = \mathbf{a}^T P \mathbf{a}$, and where the weighting matrices $\Lambda_k = \mathbf{Q}_k^{-1}$ and $\Sigma_k = \mathbf{R}_k^{-1}$.

The main differences in SLAM estimation algorithms lie in how this cost function $J(\mathcal{X}_{1:k}, \mathbf{m})$ is minimised. As noted above, the filter based approaches, EKF [56]–[58], SEIF [59], PF [60] and RBPF [61], [63] use recursive algorithms to produce the estimate of the current robot pose $\hat{\mathbf{x}}_k$. The smoothing/optimization algorithms, GraphSLAM [64], [65] and SAM [66]–[68], by contrast, operate on a batch of data using sparse least squares methods to produce the estimate of the entire state history $\hat{\mathcal{X}}_{1:k}$, where the sparse structure in the SLAM problem is exploited to make the algorithms more computationally efficient. The sparse structure arises from the fact that each landmark is only observed from a small set of poses. A key advantage of the smoothing and optimization algorithms is that they intrinsically correct all previous robot poses when new loop closures are made.

Many of the SLAM algorithms (EKF, SEIF, GraphSLAM and SAM), except those based on the PF, use a linearised form of $J(\mathcal{X}_{1:k}, \mathbf{m})$ to simplify the estimation procedure: for the filtering algorithms this linearisation enables the propagation of a Gaussian distributed state estimate. For the smoothing and optimization algorithms the linearisation enables the use of efficient, sparse least squares methods. In each algorithm, the nonlinear state and observation functions are linearised using a truncated Taylor series expansion around a linearisation point of the state, $\bar{\mathbf{x}}_k$, and the map, $\bar{\mathbf{m}}$,

$$f(\mathbf{x}_k, \mathbf{u}_k) \approx f(\bar{\mathbf{x}}_k, \mathbf{u}_k) + \mathbf{F}_k (\mathbf{x}_k - \bar{\mathbf{x}}_k), \quad (15)$$

$$h(\mathbf{x}_k, \mathbf{m}) \approx h(\bar{\mathbf{x}}_k, \bar{\mathbf{m}}) + \mathbf{H}_k (\mathbf{x}_k - \bar{\mathbf{x}}_k) + \mathbf{M}_k (\mathbf{m} - \bar{\mathbf{m}}), \quad (16)$$

where \mathbf{F}_k and \mathbf{H}_k are respectively Jacobian matrices of partial derivatives of f and h with respect to the state, \mathbf{x}_k , whilst \mathbf{M}_k is the Jacobian matrix of partial derivatives of h with respect to the map, \mathbf{m} .

An important advantage of the smoothing and optimization algorithms, related to this linearisation, is that they are able to iterate the linearised problem until convergence, whereas the filtering algorithms (EKF and SEIF) only perform a single update, which means that the filtering methods can be more prone to linearisation error.

As a consequence of linearisation error the EKF can become inconsistent [69]–[71]: a filter is consistent if the estimation error sequence, $\mathbf{x}_k - \hat{\mathbf{x}}_k$, is zero-mean and the state covariance, \mathbf{P}_k , matches the true covariance [72],

$$E[\mathbf{x}_k - \hat{\mathbf{x}}_k] = 0, \quad (17)$$

$$E[(\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T] = \mathbf{P}_k \quad (18)$$

where $E[\cdot]$ is the mathematical expectation operation. The inconsistency tends to arise in EKF-SLAM because the linearisation of f and h do not occur at the true state [71], and therefore the estimated state covariance, P_k , becomes less than the true value, i.e. the filter becomes overconfident, leading to filter divergence [69], [71]. Smoothing and optimization methods tend to be less prone to this inconsistency because the estimates are computed in a batch and are iterated to convergence [73].

The PF and RBPF methods, by contrast, avoid the linearisation of f and h : particle filters are able to fully utilise the nonlinear model and represent the non-Gaussian state posterior using a numerical sampling approach. However, in SLAM where the number of map features can be large and tends to grow without bound, sampling methods can be computationally expensive. In FastSLAM [61], the RBPF is used to exploit a factorisation to make the problem more computationally tractable, where using the product rule we re-write (1) as

$$p(\mathcal{X}_{1:k}, \mathbf{m} | \mathcal{Z}_{1:k}, \mathcal{U}_{1:k}, \mathbf{x}_0) \\ = p(\mathcal{X}_{1:k} | \mathcal{Z}_{1:k}, \mathcal{U}_{1:k}, \mathbf{x}_0) \prod_{i=1}^{n_m} p(\mathbf{l}_i | \mathcal{X}_{0:k}, \mathcal{Z}_{1:k}) \quad (19)$$

where the key insight to note is that the map features $\mathbf{m} = (\mathbf{l}_1, \dots, \mathbf{l}_{n_m})^T$ become independent when conditioned on the full robot trajectory $\mathcal{X}_{0:k}$ [61]. The robot trajectory posterior, $p(\mathcal{X}_{1:k} | \mathcal{Z}_{1:k}, \mathcal{U}_{1:k}, \mathbf{x}_0)$, is estimated using a particle filter, with n_p particles, where each particle represents a possible instance of the robot trajectory, and each particle uses n_m EKFs to separately represent and update each landmark \mathbf{l}_i in the map. This means that each EKF is low-dimensional because they only represent one landmark each, hence the naive complexity of FastSLAM is $O(n_p n_m)$, i.e. linear in the number of map landmarks. By contrast, in EKF-SLAM the full covariance corresponding to all n_m landmarks is used, meaning that updates are $O(n_m^2)$, i.e. EKF-SLAM updates have quadratic complexity. Hence, the FastSLAM algorithm is far more efficient than EKF-SLAM.

One of the problems with SLAM back-end optimization methods is that they can be sensitive to outliers arising from incorrect data associations and false-positive loop closure errors. These outliers can cause the whole SLAM system to fail because of the quadratic nature of the cost in (14), which gives undue influence to measurements with large residual errors. Therefore, robust loss functions that are less sensitive to outliers, such as the Huber loss, can be used in place of the standard quadratic loss [74]. This type of approach is used in early work in dynamic covariance scaling [75] and related methods such as switchable constraints [76], [77] and max mixtures [78], which use a tunable weighting to decrease the influence of inconsistent loop closures on the optimization. In recent work, adaptive kernels for robust cost functions have been developed [79]. There are also distinct methods that check for, and exclude, inconsistent loop closures from the optimization, such as realizing, reversing, recovering (RRR) [80].

D. VISUAL ODOMETRY AND VISUAL SLAM

Cameras have become a dominant sensor-type for robotics, owing to their versatility and usefulness across various tasks, for instance mapping and localization, structure from motion, obstacle avoidance, object detection and recognition, scene understanding and human-operator support. There is a corresponding wide and specialist literature on the use of cameras in SLAM, which therefore motivates its own section here. Cameras are also often used in *visual odometry* (VO) in robotics, which is the localization-only part of the SLAM problem, i.e. pose-estimation, without map-building [81]. The full visual SLAM (vSLAM) problem can be loosely defined as [82],

$$\text{vSLAM} = \text{VO} + \text{Global Map Optimization.}$$

where the Global Map Optimization is typically done using loop closing and smoothing/optimization algorithms discussed above.

The vSLAM/VO problem can be solved using either stereo [83]–[88] or monocular (single) cameras [89], [90]. Monocular camera systems tend to be simpler than stereo systems but have the disadvantage that they do not intrinsically perform depth perception. Therefore, monocular systems generally require multiple overlapping views of an image, from distinct perspectives, to obtain depth perception algorithmically (i.e. by triangulation). Even with algorithmic depth perception, monocular systems are still subject to scale ambiguity and scale drift.

VO algorithms are of particular interest in pipe robotics because reconstruction of the robot path from the pose, \mathbf{x}_k , intrinsically generates the pipe network map (because the robot moves within the pipe, therefore the robot pose \mathbf{x}_k defines the pipe location as well as the robot pose). Early monocular VO estimation algorithms tended to be based on filtering methods such as the EKF [89], [90], but keyframe optimization methods have become more dominant recently [88], [91]–[93]. Filtering methods marginalise

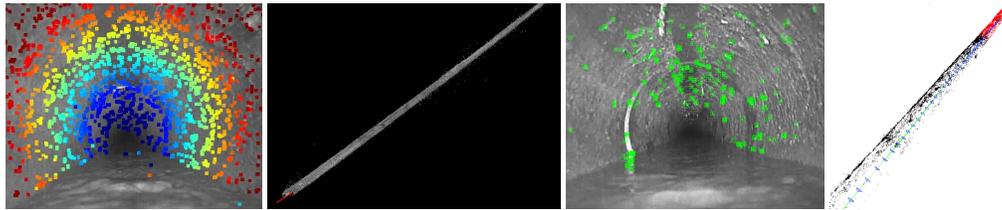


FIGURE 3. An illustration of two visual SLAM algorithms, DSO and ORB-SLAM, applied to sewer pipe images. The leftmost two images illustrate DSO, with a frame of the processed video on the left, and an estimated map of the pipe from a video sequence on the right. The rightmost two images illustrate ORB-SLAM, with a frame of the processed video on the left, and an estimated map of the pipe from a video sequence on the right.

out past poses and summarise historical information using a probability distribution. Keyframe optimization methods instead use efficient batch least squares algorithms to estimate the pose over a small number of keyframes selected from the recent frame history. One study concluded that keyframe optimization methods tend to be more accurate per unit of computing time [94]. A limitation of VO algorithms is that they do not perform loop closure in contrast to full vSLAM algorithms, so drift in pose estimates goes uncorrected.

Popular recent approaches to VO can be divided into feature-based methods, which use feature extraction to obtain image frame correspondences, such as ORB-SLAM [88], [93], and direct methods, which operate directly on pixel intensity, such as large scale direct SLAM (LSD-SLAM) [95] and direct sparse odometry (DSO) [92] (where DSO and ORB-SLAM are illustrated using our own implementations, not published, in Figure 3).

Another important class of VO method is based on deep learning, using convolutional neural networks [96], deep recurrent convolutional neural networks (DeepVO/ESP-VO) [97], [98], unsupervised deep learning (UnDeepVO) [28], generative adversarial networks [99] and deep networks driven by optic flow [100]–[102]. In [103] a deep learning VO method was developed for underwater applications, which is relevant to water distribution pipes, and showed promise compared to standard methods (although the underwater environment did prove challenging for pose estimation). Full vSLAM has also been addressed using deep learning [104]. The deep learning methods appear to give competitive results to other VO methods on benchmark problems, and have the advantage that they are end-to-end so they do not require camera calibration, feature extraction and matching, and online optimization. They do, however, tend to require large amounts of training data, which may be problematic for sewer and water pipes.

VO is often fused with an inertial measurement unit (IMU), known as visual-inertial navigation systems (VINS), or visual inertial odometry (VIO) [105], [106]. Fusing VO with an IMU tends to improve accuracy, is low cost, and for monocular systems helps to resolve the scaling ambiguity. Recent popular VIO systems include MSCKF [107], ROVIO [108], VINS-Mono [109] and Vi-DSO [110]. MSCKF is termed a loosely coupled approach and is relatively simple to

implement and computationally inexpensive (a Kalman filter fuses the VO pose estimate with the IMU, and potentially other sensors as well), whilst the others are tightly coupled (IMU data is included in the pose optimization), which tend to be more accurate [111]. The VIO problem, like many of the computer vision problems discussed in this review, has in recent years been addressed using deep learning such as in VINet [112].

E. LASER SCANNERS AND LIDAR FOR SLAM

Laser scanners and lidar (light detection and ranging) are one of the other major sensor-types, along with cameras, used in SLAM. Lidar sensors produce a scan of the environment that returns the range and bearing of nearby objects at discrete sample points - scans can be in 2D [113]–[115] or 3D [116]–[119]. The lidar SLAM problem is often divided, similarly to vision methods, into an odometry-type of problem using sequential scans for pose estimation, and separate map updating with loop closing [120].

A scan matching algorithm is typically used with lidar to estimate the pose of the robot - this is where a current scan is used with the previous scan of known pose (i.e. scan-to-scan matching), to provide an estimate of the transformation between the two scans - this transformation can be used to update the pose of the robot. However, this is essentially a type of odometry method that will drift over time. Scan matching is usually based on iterative closest point (ICP) type algorithms [113], [121], which involves a minimisation of scan matching error as a function of the transformation between scans. Scan matching error can be minimised in terms of scan points [113], or extracted features such as lines [122].

Loop closing can be performed with lidar using scan-to-map matching [123]. This includes methods based on feature extraction, which reduces computational complexity [124], histogram-based matching [51] and machine learning [125]. Submaps can also be used in lidar scan-to-submap matching [115], [126], which improves computational efficiency and enables real-time loop closing.

Lidar can also be fused with vision to overcome problems associated with the different methods, i.e. visual SLAM relies on adequate visual features to function effectively, whilst lidar can be sensitive to rapid motion (because the point cloud

can become distorted due to the robot motion interfering with the lidar scanning process) - fusion of vision and lidar can alleviate these problems [119], [127]. Vision-lidar fusion is reviewed in [128].

F. COMPARISON OF SLAM ALGORITHMS

Front-end methods for SLAM include data association and loop closing - these will have to be robust for the pipe environment because of the high likelihood of perceptual aliasing. There is a wealth of data association methods developed for mobile and manipulation robot SLAM, but data association for robot SLAM in pipes has not been addressed. For loop closing we would expect that vision-based appearance-based mapping on its own will be challenging, therefore the use of multi-sensor data fusion and prior map knowledge will be important to improve robustness.

The key advantages and disadvantages of each back-end SLAM algorithm lie in a number of factors. EKF methods are relatively simple, and tends to perform well in small to medium map problems but can be inconsistent leading to filter divergence. SEIF and FastSLAM improve on the computational efficiency of EKF-SLAM. The smoothing/optimization methods are advantageous over the filter-based methods because they treat all data in the estimation, which intrinsically leads to correction of older poses and map estimates when loops are closed, and are less prone to divergence. Hence, smoothing methods tend to be preferred in modern SLAM implementations.

SLAM in pipes will require computationally efficient solutions for relatively small robots with modest computational resources. The methods for SLAM discussed above can be extremely computationally intensive, even for the simplest methods, and especially for modern techniques aimed at robustness using multiple hypothesis methods in front-end data association [35], and back-end pose/map estimation using optimization/smoothing [129]. Therefore it is likely that SLAM in pipes will assume known maps for online localization and only perform SLAM intermittently.

V. MAPS FOR SLAM IN WATER AND SEWER PIPE NETWORKS

The choice of representation of the map of the robot's environment has implications for accuracy, precision, computational efficiency, and robustness of the SLAM system. In general, robots across different applications use a variety of descriptions of their surrounding environment, often depending on the application. In this section, the range of map representations used in the literature is described, and their usefulness for robots in pipe environments is evaluated.

There are also a number of auxiliary factors that will affect the choice of map for use for SLAM in pipes. For instance, the representation might depend to a certain extent on the locomotion and sensors used by the robot. The locomotion type can determine the space within which the robot can move, and therefore must be localised. A variety of types of robot locomotion have been developed for use in pipes, reviewed

in [130], which includes flying or swimming through a pipe with six degrees of freedom, moving along the cylindrical surface of the pipe with fewer degrees of freedom, and moving along the axis of the pipe by pressing against opposite walls giving only one degree of freedom. Therefore, the dimension of the map representation might naturally coincide with the degrees of freedom in the robot movement.

It is also worth noting that maps used by water utilities of buried pipe networks, such as that shown in Figure 1, often exist to varying degrees of accuracy. However, this type of map cannot necessarily be used directly by a robot for localization, and conversely a map estimated by a SLAM algorithm would not necessarily be of a form that would be directly useful for a human operator in a water utility company. The remainder of this section will describe different types of map used in SLAM, as opposed to maps used by humans.

A. FEATURE MAPS

Feature-based maps could be made using features at a variety of scales, and could be made using features which are generally describable such as walls or doors, or using features which depend more on the sensing mode such as notable sets of pixels in a set of camera images. These two categories of *feature*-based maps are described here.

1) POINT FEATURES

Point features can be extracted from sensor data such as images, typically corresponding to significant points in the environment which might be recognised and distinguished from other points (see Fig. 4(a)). As noted elsewhere [92], features can be sparse or dense, and direct or indirect. In this section, methods that use some level of indirect representation of points in the environment will be described, distinguished from methods using direct sensor measurements such as the distance of reflection of a lidar beam or the intensity of a pixel in a camera image.

A number of solutions exist to this problem; the Harris corner detector [131], SIFT [37], SURF [38], FAST [132], BRIEF [133], and ORB [134], being some historically popular examples. These methods show an improvement in the solution over the last two decades, with a general emphasis on efficiency for real-time application. The solutions typically detect salient points in a camera image based on pixel intensity, and describe the point using the variation in intensity of the nearby pixels. This *descriptor* can be used to find matching points in other camera images, which can be used for localization and mapping, where the point features make up the map. Similar methods could be used to extract features from data from other sensors such as sonar and lidar.

2) LARGE FEATURES

A typical robot environment might be made up of walls, doors, furniture, and people, which can be represented as geometric features. However, the variety of possible features could be much larger depending on the application, and

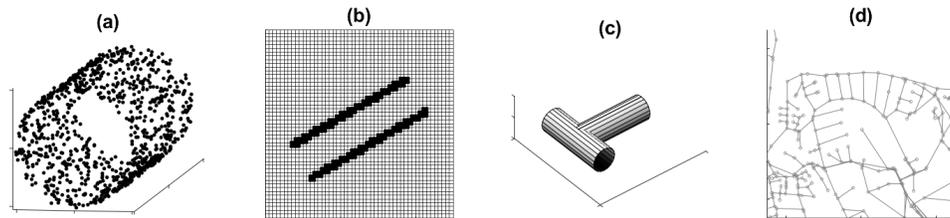


FIGURE 4. Illustrations of a number of a map types for the pipe environment. (a) A point cloud or point feature map. Each point lies approximately on a cylinder defining the pipe. (b) A two-dimensional occupancy grid representation. A three-dimensional grid would be needed for a cylindrical pipe, but would be hard to visualize. (c) A large feature map, where each feature is a cylindrical pipe, parameterized by its position, orientation, and shape. (d) A topological map of a large pipe network. Each node is a junction or manhole, and is connected to other nodes by pipes.

there would be considerable variation in features of each type suggested. Early work in SLAM used simple geometric representations such as planes, cylinders and corners [135]. Later SLAM algorithms used recognition of common features such as walls and doors [136]–[138]. An improvement to accuracy using these methods comes at the cost of reduced flexibility, and environments with unusually shaped walls and doors will be challenging for the algorithm. Knowledge of a pipe’s cylindrical shape has been used for localization in pipes (Fig. 4c) [139].

Features such as walls, or pipes in this application, can be represented in a map in a parameterised form such as B-splines. This increases the flexibility of the map representation, and has been shown to be applied in an efficient SLAM algorithm [140], [141].

B. DENSE MAPS: GRID MAPS AND POINT CLOUDS

1) GRID MAPS

A continuous metric space can be decomposed into a grid of discrete, finite sized cells. In early work on this topic these cells were relatively large and typically corresponded to notable features in the environment [142], [143]. However, works using a discrete representation of the space diverged: some becoming known as *topological maps*, which continue to use a more coarse representation of the map, and some becoming metric grid localization which use a finer grid representation [144].

In an *occupancy grid* map (Fig. 4b) [145], [146], each cell has a probability of being occupied by an object, or being empty. The occupancy probability of each cell can be updated recursively using new sensor information. Grid mapping can be performed in either 2D or 3D [147]. A key limitation of grid maps is that they require large amounts of memory if mapping over a large spatial scale at high grid resolution. Memory efficient solutions to storing grids exist using trees for both 2D [148] and 3D maps [149], [150].

The grid representation gives flexibility to the representation of the probability distribution of the position of features in the map and the position of the robot. Where an occupancy grid approach is used, the features in the map do not need to be interpreted or undergo data association, which reduces the requirements of the front-end perception module. However,

there is an inevitable loss in precision due to the discretisation of the map. In order to improve precision, grid cells need to be made small, however, as noted above this increases memory requirements.

2) POINT CLOUDS

A point cloud representation (Fig. 4) is a set of data points representing the position of features observed by the robot in the environment, typically with only position data and no further data. This might represent the observation of objects using a lidar laser scanner, or from a stereo camera. The small amount of information contained in each point means that point clouds can contain a relatively large number of individual data points. These data points are often processed as a group, so matches might be found between a new sensor scan and the existing cloud of points, for example. Point clouds have been used in a number of 3D SLAM algorithms [116], [151]. However, they do tend to be memory intensive and may not be well suited to small robots in pipes with limited memory capacity.

3) SURFACE REPRESENTATIONS

Dense surface representations acknowledge that the point features detected by cameras or lidar are part of a surface in 3D space. Surfel maps [152] (similarly named to pixel and voxel maps) have been applied to SLAM using depth cameras [153] and using 3D laser range data [154]. Truncated signed distance fields have also been applied to data from depth cameras [155].

These dense surface representations use more of the information in the sensor data, rather than extracting discrete features, which can give good performance even when sensing a surface with low texture, and the use of the surface as a concept (as opposed to point features in space) is easily applicable to the pipe environment which is made up largely of simple surfaces. However, as with point clouds, these representations are computationally expensive which is detrimental to the application to small robots in pipes.

C. TOPOLOGICAL MAPS

A topological map (Fig. 4d) is an alternative to a metric representation of the robot’s environment and state. In this case,

the map is described as a set of discrete places defined by their connectivity rather than necessarily their metric relationship. A detailed review on general topological SLAM is found in [156].

In typical mobile robot environments the environment might be discretized into a set of rooms, for example, however, the problem of discretizing the environment has a variety of possible solutions. This includes using *gateways* in the environment such as doors [157], using the *meet points* of lines equidistant from objects in the environment [158], using a square grid of cells of a fixed size [142], and separating places into *nodes* where the robot may turn and *hallways* which connect these nodes [159], however it is known that finding *distinctive places* in an environment can depend on the sensor system for a given robot [160]. Without much abstraction, a pipe network can accurately be described similarly to *nodes* and *hallways* [159], which gives a natural representation of the environment.

An advantage of a topological representation is the reduction in computational cost of localization and navigation in large environments [156]. Conversely, there is an inevitable loss of precision in the map representation, and as in many problems, a compromise must be made between cost and accuracy. Another advantage of a topological representation is that the general topological structure may be known even when precise metric information might not be available. This is especially applicable to the pipe environment, where the precise position of buried pipes may be unknown, but their connectivity can be assumed if the system is functioning as intended. Topological maps have been used in algorithms developed for localization in pipes [161], [162].

D. SEMANTIC MAPS

Semantic maps can be described as a map which contains both spatial information about the environment and classification of features, where further knowledge about these classes is available for reasoning [163]. They are therefore regarded as an enhanced map, with both geometric information and high-level qualitative features that have semantic meaning. Semantic SLAM methods are reviewed in [164].

Features with semantic meaning in pipes include:

- pipe lengths,
- T-junctions,
- Y-junctions,
- elbows,
- valves,
- pipe-joints,
- customer connections,
- fire hydrant connections, and
- manholes

These features are typical for clean and wastewater pipes and have been used as landmarks in pipe robots for many years; Infrared range sensors have been used for detecting inlets in pipes [165], laser scanners for detecting elbows [166] and T-junctions [167], and vision for detecting elbows and branches (junctions) [168], [169].

One potential advantage of using semantic maps in the pipes domain is that because existing maps held by water utilities label certain features such as manholes and valves, these features already have semantic meaning. Therefore, it is natural to include and exploit these semantic labels in prior maps of the pipe network which can then be used in semantic SLAM.

E. HYBRID MAPS

As noted above, topological maps are well suited to pipe networks and particularly the problems of path planning and exploration. Metric information is also needed, generally, to more precisely localise any defects encountered. Hybrid maps are well suited to this combination of needs, such as hybrid metric/topological SLAM [170]–[172], hybrid metric/semantic SLAM [173], [174] and hybrid semantic/topological [175]. These types of hybrid maps, combining topological and metric information, have been used successfully in mapping gas pipelines [176] and robust methods have been proposed for localization in sewer/water pipes [162].

F. COMPARISON OF MAPS

Each of the types of map described here have advantages and disadvantages in general robot localization. Typically there is a trade-off between precision and computation, and between flexibility and computation. In the application to robots in pipes compared to the general case, computation is limited significantly, flexibility is less necessary, and required precision can be variable, depending on whether the robot is trying to navigate or trying to precisely locate a fault. Therefore an effective map representation for the pipe environment will likely be a hybrid representation, using the environment topology to an extent for efficient path planning and exploration, and a metric representation to precisely map the network and locate faults.

VI. SENSORS FOR MAPPING AND LOCALIZATION IN WATER AND SEWER PIPES

Sensors are a key factor to consider when designing a SLAM system. The back-end algorithms tend to be common and interchangeable across domains, but the sensors need to be selected to suit the environment. This section reviews different sensors that have been used for robot navigation in pipes.

A. INERTIAL AND ODOMETRY DEAD RECKONING SENSORS WITH DRIFT CORRECTION

The most simple methods of in-pipe localization for robots have been based on dead-reckoning techniques, usually combining an inertial measurement unit (IMU) with some form of odometry [177], [178]. Odometry sensing has been used with tether systems [179], which limits the distance the robot is able to travel, and alternatively on-board wheel odometry for untethered robots [180], which is less restrictive. The accuracy of IMU-odometry localization systems has recently been analysed in [181], demonstrating that errors using a high-grade (quasi-tactical) IMU were not more than 0.25% and

0.1% of the pipe lengths in the horizontal and vertical directions respectively (tested pipe lengths ranged from 100 m to 1700 m). Micro-electromechanical system (MEMS) IMUs have also been proposed for use in small diameter pipes, because of their reduced size, which might require improved algorithms to compensate for their lower accuracy (e.g. use of the cubature Kalman filter) [182].

Drift is the key problem with dead-reckoning sensors, and so dead-reckoning has also been combined with drift correction methods using detected landmark locations. The landmarks can either be naturally occurring, such as pipe-length joints [180], [183]–[185] or deliberately added to the pipe network, such as GPS-located above ground reference stations [186], [187]. Pipe-length joints can be detected using magnetic flux leakage and electromagnetic acoustic transducers [180], [184], or from the vibration signal from an accelerometer as the robot passes through the joint [183].

The use of naturally occurring landmarks for drift correction, such as pipe-length joints, adds minimal cost to the localization solution (i.e. just the cost of the sensors), but is only applicable to scenarios where the landmarks have some *a priori* known position - e.g. this is the case in certain pipelines where each pipe length is known with certainty. This type of approach has been researched in water distribution pipes, combined with IMU dead-reckoning [185] and is promising because it is low-cost and provides the needed drift correction to dead reckoning methods. The disadvantage with deliberately adding reference landmarks to a pipe network is that it increases the cost and could become prohibitively expensive when considering the hundreds of thousands of kilometres of existing pipe networks.

Sensor motes for water distribution pipes have been proposed as an alternative to conventional robots, that could use an IMU plus additional sensing for drift correction in localization [188]. The sensor mote is passive and carried by flow through the pipe network. The key advantages of these simple sensor motes, over more sophisticated robots, is that they are likely to be more easily miniaturised, be cheaper, consume less power and be more robust. In the context of small diameter water pipe environments, and low overheads in the water utilities industry preventing uptake of expensive and complex technology, these benefits are appealing.

B. CAMERAS

Cameras are very often included on pipe inspection robots so that damage can be detected by visual inspection, such as in MAKRO [189], KANTARO [190], MRINSPECT [191], PipeTron [192] and EXPLORER [193]. Therefore, cameras and visual information about the surroundings are a natural choice to pursue for navigation. Early work used vision to estimate distance travelled along the pipe only, via an image mosaicking algorithm and a laser range finder for depth perception [194]. However, modern VO systems with keyframe optimization of the type described above tend to be used now [139], [195], [196].

Fisheye cameras are most often used for VO in pipes because of the narrow structure of the environment [139], [196]–[199] (although a panoramic camera has also been tested [200]). In a pipe, when the camera is panned along the pipe axis, the distant surface of the pipe in the image suffers from projective deformations, whilst the near pipe surface appears in peripheral regions of the images clearly. The distant features all have a low parallax angle which degrade triangulation in VO. Hence, use of fisheye cameras tends to lead to more accurate VO estimates in pipes.

Stereo cameras have advantages over monocular systems because they have a fixed baseline between the two cameras, which aids triangulation, and they also automatically resolve scale ambiguity. In [195], two cameras were aligned facing upwards toward the pipe surface, which enabled a large amount of stereo image overlap for depth perception. By contrast, in [201], an axial-stereo vision system was proposed, where the cameras faced forward along the pipe [201]. Simpler, monocular systems have also been used in pipes, such as in [139], [196], [202]. In [139] structured lighting was used to recover the scale factor in monocular VO by minimising the reprojection error of two laser spots [139]. Distinct from both stereo and monocular systems are depth (RGB-D) cameras, which have the advantage that they can also be readily used to detect obstacles in the pipe [203]–[205].

Many VO algorithms for pipes leverage cylindrical information which exploits prior knowledge of the shape of the environment to improve pose estimation accuracy, and to help resolve scale ambiguity for monocular systems [196], [202], [206].

Robust methods have been developed that optimize the map points in the optimization by enforcing cylindrical regularity [139], [207]. Cylindrical regularity has also been used as prior knowledge for local pose optimization, to further improve accuracy [199]. One issue with using cylindrical information, however, is that it is necessary to detect features from the cylinder surface to match against a cylinder model. This has not been well researched because most papers assume clean, empty pipes - perhaps ideal for gas pipes, but this would not necessarily be the case in sewers (and even water pipes can have non-clean surfaces where biofilms accrue [208]). One possible solution to this problem is a method for detecting outliers proposed for 3D occupancy grid maps [209], but this is computationally intensive, so not necessarily well suited to real-time implementation.

Cameras can also be used in appearance-based SLAM methods [36], [43] and to recognise landmark features in pipes such as T-junctions, elbows etc. [168], [169], [176], [210]. In sewer pipes, VO has been combined with manhole recognition to correct drift, which is an appealing approach [203], [204]. Also, landmark recognition can be used to construct a topological pipe network map to enable the application of topological SLAM, and efficient topological path planning methods [211], [212].

C. LASER SCANNERS

Laser scanners have been used for many years for inspection in pipes [17], [213] and also for navigation in pipes including on the robots KANTARO [214] and MRINSPECT [215], [216].

The laser scanners on sewer robots have typically been used to recognise landmark features in sewer pipes, such as T-junctions, joints and elbows [167], [214], [216]. In KANTARO, the method of landmark detection is based on using the range pattern obtained from the laser to classify different types of landmark.

Laser scanners have also been used with cameras to improve the detection of landmarks - in one version of KANTARO, a stereo camera system first computes distance to a captured image and then the laser scanner is used to classify the landmark as a manhole or joint [217]. It was also noted in [217] that conventional laser scanners were too bulky at that time for sewer inspection and so they designed a custom laser-scanner more suited to the sewer pipe environment.

D. ACOUSTIC AND RADIO FREQUENCY EMITTER-RECEIVER SENSING

A number of emitter-receiver methods have emerged recently that seek to overcome the potential limitations of vision using acoustic and radio frequency sensing.

High frequency acoustic sensing (an ultrasonic sensor) [218], mid-frequency acoustic sensing (a hydrophone sensor) [219]–[221], and radio frequency (RF) sensing [222], [223] have all been applied to localization in pipes with the similarity that they all using an emitter-receiver approach to create a type of 1D spatial map along the pipe that is continuous, and hence more feature-rich than intermittent visual landmarks.

Low-frequency acoustic sensing (a speaker and microphone) [224], [225] has been used to make absolute measurements of the position of the robot in the pipe, relative to either a fixed source in [224], and using acoustic echoes in [225]. In either case, the absolute measurements of position have an advantage over visual odometry, which will drift over time.

The robot localization methods developed in [218]–[221], [225] have the advantage that the emitter-receiver unit is carried on-board the robot, whilst in contrast, the other acoustic [224] and RF [222] localization methods require the emitter to be placed in a fixed location in the pipe, with the receiver on-board the robot.

The ultrasonic sensing method developed in [218] is suited to plastic water pipes because it uses an ultrasonic transceiver to sense terrain profile through the plastic pipe wall. In contrast, the sensing methods developed in [219]–[221] use a hydrophone to excite pipe vibration in metal pipes. The initial work [219], [220] used an EKF and particle filtering methods for robot localization, whilst [221] developed a GraphSLAM method for localization.

E. ABOVE-GROUND SENSING METHODS INCLUDING GROUND PENETRATING RADAR

Above ground methods are often used for pipe detection and localization. This section briefly surveys some of these methods because they can form a useful part of an in-pipe robot SLAM system, by providing prior information and for fusing with in-pipe map estimates to improve asset mapping.

Ground penetrating radar (GPR) uses electromagnetic waves (typically in the MHz range) to detect and locate below ground targets [9]. GPR has long been proposed as a tool that can perform subsurface detection of buried utility infrastructure, including pipes [226]. More recent reviews on the use of GPR for locating underground utilities are given in [10] and [227]. GPR is often combined with additional sensors to more effectively locate targets, e.g. GPS [228], electric fields [229] and cameras [230], [231], and also multiple sensor suites such as GPR, GPS, lidar, cameras and encoders [232].

GPR data can also be combined with data from utility company records [233], although such data might not always be available or accurate. Various data fusion algorithms have also been developed to fuse GPR with other data sources based on Bayesian methods [234], [235] and Dempster-Schafer methods [236], where the latter avoid the need of the Bayesian methods to specify a probability distribution for each data source.

Additional above ground methods have been proposed for locating buried infrastructure, for instance in [237] where multiple above ground sensors are fused including GPR, Passive Magnetic Fields (PMF), Magnetic Gradiometer (MG), Low Frequency Electromagnetic Fields (LFEM) and Vibro-Acoustics (VA). Any of these techniques have potential for fusing with in-pipe robotics to improve detection accuracy.

F. COMPARISON OF SENSORS

The different types of sensor have various advantages and disadvantages for SLAM in pipes (summarised in Table 5). The basic inertial and odometry dead reckoning sensors provide simple localization but are subject to drift, so need to be used with landmark recognition methods to correct drift.

Cameras are one of the most widely used and mature localization and mapping methods in robotics, and have been used successfully in pipes. However, cameras do have certain disadvantages for use in pipes: the environment might lack visual features and so present problems of perceptual aliasing; vision-processing is usually computationally intensive so can require sophisticated and large computational hardware (which can be power intensive); the camera lens might become dirty and occluded by objects, particularly in sewer pipes; and the pipe environment is dark, so a light source is required (which consumes battery power).

Laser scanners/lidar are another popular type of sensor for SLAM, but have only been used, it would appear, for landmark recognition of e.g. manholes and elbow bends, not for pose estimation along the pipe lengths (this may be due to

TABLE 5. Advantages and disadvantages of different sensor-types for localization and SLAM in pipes.

Sensor-type	Advantages	Disadvantages	Key papers
Inertial measurement units (IMUs) and odometry	Simple, cheap, effective; impervious to pipe environment challenges such as darkness, feature sparsity and perceptual aliasing	Drifts over time so needs drift correction using known landmark locations such as in-pipe joints or above-ground reference stations	IMU+odometry [179], [180], [181]; drift correction using pipe joints [180], [183]–[185], and above-ground reference stations [186], [187]
Cameras	Effective for localization along pipe lengths using visual odometry, as well as loop closing at landmarks such as manholes and junctions using place recognition methods	Not well suited to pipes because of the dark environment and small parallax angle between frames; also may suffer from failure due to lack of visual features and perceptual aliasing due to high similarity in visual appearance between locations	Monocular VO [196], [202]; Stereo VO [195]; VO+cylindrical regularity [139]; vSLAM+cylindrical regularity [206]; RGB-D cameras [203], [204]
Laser scanners/lidar	Effective for landmark recognition such as manholes, junctions and elbows	Not well suited to localization along pipe lengths due to lack of features with distinct range profiles	Landmark recognition [214], [216], [167]
Acoustics and radio	Effective for absolute, drift-free ranging and impervious to dark environment	New and emerging technology; 1D distance estimation along the pipe-length only; some methods require emitter to be placed in fixed location; can be challenging to data associate acoustic echo signals with landmarks	Ultrasonic [218]; hydrophone [219]–[221]; RF [222], [223]; low frequency acoustic (fixed emitter) [224]; low frequency acoustic-echo (on-board emitter/receiver) [225]
Above-ground sensors including ground penetrating radar (GPR)	Can receive GPS signals for absolute, drift-free localization; relatively simple to fuse with geographic information system (GIS) prior data	Requires above-ground systems, which raises a number of issues to do with interference with humans in unstructured environments	GPR+camera [230]–[232]; Fusion with GIS data [234], [235]; multisensor data fusion [237]

the fact that scan-to-scan pose estimation does not work well in pipe lengths because they tend to lack useful discriminative features for range-based pose estimation). Lidar is advantageous for pipes because it does not require a light source but has the drawbacks that the sensors tends to be more costly and bulky than cameras.

Acoustic sensing is appealing for pipes because acoustic waves tend to propagate well in this environment, and they do not require a light source. However, some acoustic methods require a fixed sound source, which limits mobility.

Above-ground methods are potentially useful where available but they would appear to be more complex to implement autonomously due to the fact they would need to operate in the unstructured above-ground environment.

Therefore, it would appear that multi-sensor data fusion is the most appealingly approach for sensor choice for SLAM in water and sewer pipes. Visual odometry, acoustic methods and inertial sensing appear well suited to localization along pipe lengths. Drift correction methods are critical for odometry-type methods, using landmarks such as manholes or potentially pipe joints if they occur at predictable locations. Vision and laser scanners appear well suited to the problem of landmark recognition at manholes, elbows and junctions. A publicly available dataset incorporating multiple sensors for localization in sewer tunnels is described in [238], including inertial, camera and laser sensors, which will enable the interested reader to investigate these sensors for themselves.

VII. FUTURE CHALLENGES

A. SINGLE ROBOT SLAM IN PIPES

The methods discussed so far for SLAM in water and sewer pipes have mainly been tested in lab environments or

small scale outdoor experiments. To bridge the gap to real-world use, a number of future challenges must be overcome. Cadena *et al.* [22], in one of the most recent and comprehensive reviews of modern SLAM, highlights the following areas as future general areas to address in SLAM: robust performance, high-level understanding, resource awareness, and task-driven inference. These are particularly true in the domain of pipes and aspects of these are considered below.

1) LONG TERM ROBUST AUTONOMOUS OPERATION AND SCALABILITY

SLAM for water and sewer pipe networks must be robust in the long term. It would be unacceptable for the water industry to take up a robotics inspection system that becomes lost in the pipe network, adding to the problem of faults rather than solving the problem. The review in Cadena *et al.* [22] concludes that long term robust SLAM is not currently possible, and much fundamental work remains to be done to solve this problem. Recent work has addressed developing more robust SLAM back-end optimization algorithms using multiple hypothesis methods in MH-iSAM2 (multiple hypothesis iSAM2) [129]. However, guaranteed fail-safe SLAM is still not a solved problem.

So the question remains on how to leverage current SLAM technology for robust operation in the domain of pipes in the near future. One practical way to address this problem is to modify the pipe network with locating beacons. This approach bears similarity to the idea of using above-ground reference stations mentioned above [186], [187] but modified for buried water and sewer pipes networks. There is a rising interest in developing smart monitoring for infrastructure, for instance smart pipes [239] and smart manholes [240], with

RFID tags that communicate with the cloud, which can be used for autonomous pipeline monitoring [241], [242]. These devices could be used intermittently through a pipe network to provide recovery points for re-localization if the SLAM system fails. Such modifications of the pipe network might be costly, but this concern is alleviated if we consider that some amount of modification of the pipe network would be necessary anyway to provide communication hubs for transmitting inspection data.

Long term operation in pipes also raises the issue of scalability, where data storage and processing needs will continue to grow without bound. To address the problem of scalability, if we assume that pipe inspection robots are small, low-powered devices and that the pipe network map changes very slowly over time, an appealing strategy would be to only update the map at communication hubs by transmitting data to the cloud. This would mean robots perform localization-only whilst travelling the pipe network, in-between map updates whilst stationary at hubs, alleviating the computational burden of performing full SLAM during real-time navigation. This synergy between updating the map at communication hubs, possibly via the cloud, and robustly re-localising at hubs with full certainty would provide a natural solution to fail-safe, scalable SLAM in pipe networks. The challenge becomes one of minimising the cost of modifying the pipe network with smart communication hubs and ensuring sufficient coverage to guarantee fail-safe operation of the SLAM system in between returns to a hub. This implies that both the SLAM method and the smart pipe infrastructure should be co-designed and evaluated in a single complete framework in order to minimise costs and maximise robustness.

2) FAILURE-AWARE AND TRUSTWORTHY ROBOTICS

A major emerging issue across robotics and AI is the development of trustworthy solutions [243]–[248]. In the domain of water and sewer pipes, inspection robotics will provide key information on whether there is a fault, the type of fault, the severity, and also the location. In turn this will lead to human operators making decisions on *whether* to effect a repair and *where* to do it if so. Decisions to repair buried pipes will lead to excavations that are costly and disruptive; if the robot system makes errors this will severely impact the uptake of the technology (and potentially harm the uptake of robotics for many years).

Therefore, developing trustworthy robotic solutions for pipe networks is a key challenge for the future. This requires solutions able to work under different environmental changes and provide some characterisation of the level of trust in the solutions. It will be important to know under what conditions the approaches are fully trustful - for instance if a camera has become partially obscured, leading to inaccuracies in a visual SLAM system, the robot should be aware of this and be able to communicate the information. In general, the robot system should be *failure-aware*. This means, in the context of SLAM, that the robot should be able to intelligently detect failure of

the SLAM system, if it occurs, and be able to communicate this to human operators.

Although currently a huge effort is focused on the development of reliable and trustworthy methods, the development of trustworthy solutions for intelligent robot systems, especially in pipe networks, is still lacking.

3) PATH PLANNING AND ACTIVE SLAM

The main tasks of a water/sewer pipe inspection robot are to detect and locate faults, and also, as outlined here, map the pipe network. In order to do this, the robot must explore the pipe network thoroughly to map it, and also continue to traverse the network in order to ensure complete and consistent coverage for long-term fault detection [249]. These primary goals of the robot link naturally to an *active SLAM* framework [250], which refers to the full problem of path planning in the context of the mission, as well as mapping and localization.

Active SLAM has been addressed using many of the main frameworks for back-end pose and map estimation including EKF-SLAM [250], particle filtering SLAM [251] and smoothing/optimization SLAM [252], [253]. The typical formulation of the active SLAM problem consists of three steps [251]:

- 1) The robot identifies possible targets to visit: *frontier targets*, from the boundary of the explored region of the robot's map, where the boundary can be identified by regions where the map certainty drops below a threshold; *trajectory targets*, where map certainty is high, which the robot can return to in order to reduce location uncertainty.
- 2) The robot predicts the expected information gain associated with the visiting a frontier target, along with the risk of incorrect trajectory approximation.
- 3) The robot travels to the target and then decides if it is necessary to continue or terminate the task.

Active SLAM has not been addressed in the context of buried pipe networks but would appear to be an appealing strategy for this domain. One of the main challenges still to address is the prediction of the expected information gain [22], which can be computationally intensive and so not well suited to small, low-powered pipe robots. Therefore, computationally efficient solutions need to be found for this problem. In addition, active SLAM in water/sewer pipes will be complicated by the fact that it is generally easier and more energy efficient to move in the direction of flow, which might nuance solutions for this environment.

B. PRIOR MAP GENERATION FOR SLAM IN WATER AND SEWER PIPES

SLAM methods are often developed with the assumption that there is zero prior knowledge of the environment. However, this is far from the case for water and sewer pipe networks. Prior information of pipe networks maps can come from a variety of sources. A challenge for the future is to leverage

these data sources, fuse them together and synthesise prior maps for robot navigation and SLAM.

Geographical information systems (GIS) have been used for many years in the water industry to capture, store, manipulate, analyze, and display spatial information for water and sewer pipe networks [254]–[256]. GIS pertaining to water/sewer pipe networks tend to include pipe locations and access points such as manholes and fire hydrants. Manhole recognition from inside the sewer pipe, combined with GIS map data, has been used to correct drift in robot localization [204]. However, recorded locations of pipes and access points in GIS maps might differ from what is actually present on the ground due to errors in data collection and data entry [255]. Therefore, the use of GIS alone would seem insufficient for the task of synthesising prior maps for SLAM.

There is a growing literature on mapping above-ground infrastructure using observations gathered from a variety of sensors including GPS, in conjunction with object recognition using machine learning. In particular, manholes can be mapped from above ground, and can also be detected below ground, within the pipe [257], hence make ideal landmarks. The problem of automatically mapping manholes from above ground has been studied using cameras [258] and laser scanners [259]–[261], as well as from the air using cameras in unmanned aerial vehicles (UAVs) [262]. More recently, DCNNs have been applied to the task of object detection and recognition for manholes [263], [264], including localising manholes [265]. Manhole cover detection with DCNNs has also been used to reconstruct the likely underground pipe locations between manholes using industry rules [266], [267].

The availability of GIS maps and data-driven maps of above-ground infrastructure raises the challenge of fusing these sources to generate a map prior. In [268] the authors use GIS data to improve the localization and detection of infrastructure objects in camera images in conjunction with a standard object detection algorithm [269]. The problem of fusing spatial data, e.g. GIS data, with images has also been studied in [270]. The fusion of different objects from different data sets (GIS and data-driven) also raises the challenge of handling uncertainty. Recent attempts have been made to quantify uncertainty in object detection from images with deep neural networks [271] but in general this is not a solved problem. Bayesian data fusion has been used to fuse data both in 2D in [234], where the focus was on GPR and GIS data, which has more recently been extended to Bayesian fusion in 3D using a wide variety of sensors [235].

There are not many studies currently on how to effectively incorporate priors into a SLAM system. In [272], a framework was developed for using Bayesian priors for SLAM in buildings. One issue highlighted was the care needed in tuning the certainty associated with the map prior: too much certainty leads to the map prior dominating the SLAM system, even when the robot senses discrepancies in live operation, and vice-versa. There is an opportunity, therefore, to generalise the results from [272] to form a framework for using prior knowledge in SLAM for pipe networks.

C. MULTI-ROBOT SLAM IN WATER AND SEWER PIPES

Cities of the future will likely have teams of robots cooperatively mapping water and sewer pipes. This will be essential to ensure full coverage of the pipe network. Multi-robot SLAM in water and sewer pipes has not yet been addressed in the literature, and raises a number of problems beyond single robot SLAM, regarding both the data and mapping-localization algorithms, which are covered in this section.

Key questions on data include [273]:

- 1) What type of data will be shared? Will it be raw or processed data?
- 2) Considering issues of limited coverage and bandwidth, how will data be shared?
- 3) Where and how will data be processed? With centralised, decentralised, distributed, or hybrid architectures and methods?
- 4) Are the methods scalable and applicable to large pipe networks? How do we deal with missing data and data transmitting at different sampling rates?

Regarding processing of the data, the three main approaches to multi-robot data fusion are centralised, distributed and decentralised [274], [275]:

- 1) Centralised data fusion: raw sensor data from multiple sensors are fused in a centralised processing node to produce a state estimate.
- 2) Distributed data fusion: raw sensor data is processed locally to produce a state estimate in each sensor, then the multiple state estimates are fused in a centralised processing node. This could be either:
 - a) Fusion under known data correlation.
 - b) Fusion under unknown data correlation. This requires additional estimation.
- 3) Decentralised data fusion: raw sensor data is processed locally and fused locally at each node.

These questions are highly relevant to the inspection of real pipe networks due to potentially limited communication range, bandwidth, memory and processing power for small robots in pipes.

The centralised fusion approach is theoretically optimal but may not be ideal in practice because it requires large communication bandwidth, large processing capability in the central node, and is not robust due to the possibility of central node failure. However, if the cloud is used to perform the centralised fusion then this can be considered robust to failure, and we could assume that it would also have sufficient bandwidth and processing capability, and is appealing if robots have to return to communication hubs to transmit data on mapping and fault detection above-ground.

The algorithms for multi-robot SLAM tend to be extensions of the single robot SLAM algorithms described above in section IV, such as filtering methods based on the EKF [276], the extended information filter [277], [278], and the particle filter [279], and smoothing/optimization methods based on GraphSLAM [280], SAM [281], [282] and iSAM2 [283]. There are also a number of works that specifically address

multi-robot visual SLAM using both centralised [284], [285] and decentralised/distributed fusion [286], [287], where the approach in [285] specifically targets the use of the cloud to perform centralised processing for the computationally intensive map optimization and data storage - an appealing approach for low-powered robots in pipes.

A key part of multi-robot SLAM is merging maps from all robots to construct a single, global map of the environment. To merge maps, generally either the initial poses of the robots must be known, or the robots must rendezvous to ascertain each other's pose, or the maps must overlap [273]. Robust methods for selecting consistent measurements for map merging now exist to solve this problem [288]. Map merging is likely to be simplified in the pipe environment because it would be feasible to obtain the initial pose of each robot by taking GPS readings of the robots at the point of entry to the pipe network, or exploiting prior knowledge of the location of the pipe access points - this would significantly simplify multi-robot SLAM in pipes. Robots could also potentially rendezvous in pipes to communicate and share map data, which would also serve to provide line-of-sight pose estimation. Alternatively, the robots could rendezvous with communication hubs throughout the pipe network and transmit their own estimated pose for centralised data-fusion and SLAM.

Finally, it is worth noting that if the initial pose of the robot is known in the world coordinate frame (from the point of entry into the pipe network), and features with known locations in the above-ground world coordinate frame are used in SLAM, such as manholes as in [203], [204] or fire hydrants, then the mapping can be directly performed by each robot in the world coordinate frame, which should greatly simplify the map merging process. The main problem would therefore be the uncertainty around multiple robots recognising common landmarks in the pipe network on which to base map merging, i.e. the data association problem. Robust methods would need developing to handle this problem, possibly based on multi-hypothesis SLAM.

VIII. CONCLUSION

This paper has presented a review of SLAM for inspection robots operating in buried water and sewer pipes. The review focused initially on the motivation that the buried pipe networks represent huge current and future investment and are an ideal domain where autonomous robots can make an important impact in future smart cities.

We reviewed the requirements of a robot system, focusing on mapping the buried assets and locating defects. We reviewed the main SLAM methods used in robotics currently (where smoothing/optimization methods tend to dominate over EKF and particle filtering methods popularised in the early days of SLAM), and brought recent reviews up to date with a discussion of the recent impact of deep learning in loop closing and visual odometry. We considered different map-types used in SLAM and concluded that hybrid methods, e.g. metric/topological, metric/semantic, and

semantic/topological show promise for this environment, where topological maps are useful over a large spatial scale for efficient path planning, whilst metric information is needed for e.g. localising defects. We reviewed the various sensors-types used in SLAM and found that a wide range of sensors have been successfully used in pipes, implying that multi-sensor data fusion is probably the most appealing approach to maximise robustness.

And finally, we looked at future challenges focusing on single robots, with challenges of long term robust and scalable operation, trustworthy robotics, and active SLAM; the development of prior maps using existing data with above ground landmark mapping; and SLAM in pipes with multi-robot teams.

REFERENCES

- [1] P. Rizzo, "Water and wastewater pipe nondestructive evaluation and health monitoring: A review," *Adv. Civil Eng.*, vol. 2010, pp. 1–10, May 2010.
- [2] GOV.UK. (Accessed: Jul. 17, 2020). *Water and Treated Water*. Accessed: 2015. [Online]. Available: <https://www.gov.uk/government/publications/water-and-treated-water/water-and-treated-water>
- [3] Bipartisan Policy Center Executive Council on Infrastructure Water Task Force. (Accessed: Sep. 29, 2020). *Understanding America's Water and Wastewater Challenges*. Accessed: 2017. [Online]. Available: <https://bipartisanpolicy.org/wp-content/uploads/2019/03/BPC-Infrastructure-Understanding-Americas-Water-and-Wastewater-Challenges.pdf>
- [4] (Accessed: Aug. 18, 2020). *Government Office for Science. UK Water Research and Innovation Framework: Taking Responsibility for Water*. Accessed: 2012. [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/294176/11-1416-taking-responsibility-for-water-summary.pdf
- [5] U.S. Environmental Protection Agency. (Accessed: Sep. 29, 2020). *Clean Watersheds Needs Survey 2012 Report to Congress EPA-830-R-15005*. Accessed: 2016. https://www.epa.gov/sites/production/files/2015-12/documents/cwns_2012_report_to_congress-508-opt.pdf
- [6] The Consumer Council for Water. (Accessed: Jul. 20, 2020). *Water, Water Everywhere? Delivering Resilient Water and Waste Water Services Everywhere (2017-2018)*. Accessed: 2017. [Online]. Available: <https://www.ccwater.org.uk/research/water-water-everywhere-delivering-resilient-water-and-wastewater-services-2017-18/>
- [7] American Society of Civil Engineers. (Accessed: Sep. 29, 2020). *Failure to Act: The Economic Impact of Current Investment Trends in Water and Wastewater Treatment Infrastructure*. Accessed: 2011. [Online]. Available: http://www.asce.org/uploadedfiles/issues_and_advocacy/our_initiatives/infrastructure/content_pieces/failure-to-act-waterwastewater-report.pdf
- [8] R. Wirahadikusumah, D. M. Abraham, T. Iseley, and R. K. Prasanth, "Assessment technologies for sewer system rehabilitation," *Autom. Construct.*, vol. 7, no. 4, pp. 259–270, May 1998.
- [9] D. J. Daniels, Ed., *Ground Penetrating Radar*, 2nd ed. London, U.K.: IET, 2004.
- [10] C. Plati and X. Dérobert, *Inspection Procedures for Effective GPR Sensing and Mapping of Underground Utilities and Voids, With a Focus to Urban Areas*. Cham, Switzerland: Springer, 2015, pp. 125–145.
- [11] S. B. Costello, D. N. Chapman, C. D. F. Rogers, and N. Metje, "Underground asset location and condition assessment technologies," *Tunnelling Underground Space Technol.*, vol. 22, nos. 5–6, pp. 524–542, Sep. 2007.
- [12] British Standards Institution. (Accessed: Aug. 29, 2020). *Pas128:2014 Specification for Underground Utility Detection, Verification and Location*. Accessed: 2020. [Online]. Available: <https://shop.bsigroup.com/ProductDetail/?pid=00000000030267400>
- [13] WRC Infrastructure. (Accessed: Aug. 29, 2020). *Sahara Technology*. Accessed: 2020. [Online]. Available: <https://www.wrc-infrastructure.co.uk/technologies/sahara%C2%AE/>

- [14] WRC Infrastructure. (Accessed: Aug. 29, 2020). *SmartBall Technology*. Accessed: 2020. [Online]. Available: <https://www.wrc-infrastructure.co.uk/technologies/smartball%C2%AE/>
- [15] J. M. M. Tur and W. Garthwaite, "Robotic devices for water main in-pipe inspection: A survey," *J. Field Robot.*, vol. 27, no. 4, pp. 491–508, Jul. 2010.
- [16] Z. Liu and Y. Kleiner, "State of the art review of inspection technologies for condition assessment of water pipes," *Measurement*, vol. 46, no. 1, pp. 1–15, 2013.
- [17] O. Duran, K. Althoefer, and L. D. Seneviratne, "State of the art in sensor technologies for sewer inspection," *IEEE Sensors J.*, vol. 2, no. 2, pp. 73–81, Apr. 2002.
- [18] L. Guan, Y. Gao, H. Liu, W. An, and A. Noureldin, "A review on small-diameter pipeline inspection gauge localization techniques: Problems, methods and challenges," in *Proc. Int. Conf. Commun., Signal Process., their Appl. (ICCSA)*, Mar. 2019, pp. 1–6.
- [19] N. M. Yatim, R. L. A. Shauri, and N. Buniyamin, "Automated mapping for underground pipelines: An overview," in *Proc. 2nd Int. Conf. Electr., Electron. Syst. Eng. (ICEESE)*, Dec. 2014, pp. 77–82.
- [20] H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping: Part I," *IEEE Robot. Autom. Mag.*, vol. 13, no. 2, pp. 99–108, Jun. 2006.
- [21] T. Bailey and H. Durrant-Whyte, "Simultaneous localization and mapping (SLAM): Part II," *IEEE Robot. Autom. Mag.*, vol. 13, no. 3, pp. 108–117, Sep. 2006.
- [22] C. Cadena, L. Carlone, H. Carrillo, Y. Latif, D. Scaramuzza, J. Neira, I. Reid, and J. J. Leonard, "Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age," *IEEE Trans. Robot.*, vol. 32, no. 6, pp. 1309–1332, Dec. 2016.
- [23] OFWAT. (Accessed: May 3, 2020). *PR19 Final Determinations: Overview of Final Determinations*. Accessed: 2019. [Online]. Available: <https://www.ofwat.gov.uk/publication/pr19-final-determinations-overview-of-final-determinations/>
- [24] American Society of Civil Engineers. *Standard Guidelines for the Collection and Depiction of Existing Subsurface Utility Data*. Reston, VA, USA: American Society of Civil Engineers, 2002.
- [25] A. C. Twort, D. D. Ratnayaka, and M. J. Brandt, *Water Supply*, 6th ed. Oxford, U.K.: Butterworth-Heinemann, 2009.
- [26] D. Butler, C. Dignan, C. Makropoulos, and J. W. Davies, *Urban Drainage*, 4th ed. Boca Raton, FL, USA: CRC Press, 2018.
- [27] A. Kendall, M. Grimes, and R. Cipolla, "PoseNet: A convolutional network for real-time 6-DOF camera relocalization," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 2938–2946.
- [28] R. Li, S. Wang, Z. Long, and D. Gu, "UnDeepVO: Monocular visual odometry through unsupervised deep learning," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 7286–7291.
- [29] K. Sachedina, A. Mohany, and U. of Ontario, "A review of pipeline monitoring and periodic inspection methods," *Pipeline Sci. Technol.*, vol. 2, no. 3, pp. 187–201, Sep. 2018.
- [30] Y. Bar-Shalom, F. Daum, and J. Huang, "The probabilistic data association filter," *IEEE Control Syst.*, vol. 29, no. 6, pp. 82–100, Dec. 2009.
- [31] J. Neira and J. D. Tardós, "Data association in stochastic mapping using the joint compatibility test," *IEEE Trans. Robot. Autom.*, vol. 17, no. 6, pp. 890–897, Dec. 2001.
- [32] H. J. S. Feder, J. J. Leonard, and C. M. Smith, "Adaptive mobile robot navigation and mapping," *Int. J. Robot. Res.*, vol. 18, no. 7, pp. 650–668, 1999.
- [33] D. Z. Wang, I. Posner, and P. Newman, "Model-free detection and tracking of dynamic objects with 2D lidar," *Int. J. Robot. Res.*, vol. 34, no. 7, pp. 1039–1063, Jun. 2015.
- [34] I. J. Cox and J. J. Leonard, "Probabilistic data association for dynamic world modeling: A multiple hypothesis approach," in *Proc. 5th Int. Conf. Adv. Robot. Robots Unstructured Environ.*, 1991, pp. 1287–1294.
- [35] J. Wang and B. Englot, "Robust exploration with multiple hypothesis data association," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2018, pp. 3537–3544.
- [36] M. Cummins and P. Newman, "FAB-MAP: Probabilistic localization and mapping in the space of appearance," *Int. J. Robot. Res.*, vol. 27, no. 6, pp. 647–665, 2008.
- [37] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 2, Sep. 1999, pp. 1150–1157.
- [38] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded up robust features," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2006, pp. 404–417.
- [39] M. Cummins and P. Newman, "Appearance-only SLAM at large scale with FAB-MAP 2.0," *Int. J. Robot. Res.*, vol. 30, no. 9, pp. 1100–1123, Jun. 2011.
- [40] D. Gálvez-López and J. D. Tardos, "Bags of binary words for fast place recognition in image sequences," *IEEE Trans. Robot.*, vol. 28, no. 5, pp. 1188–1197, Oct. 2012.
- [41] M. J. Milford and G. F. Wyeth, "SeqSLAM: Visual route-based navigation for sunny summer days and stormy winter nights," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2012, pp. 1643–1649.
- [42] S. M. Siam and H. Zhang, "Fast-SeqSLAM: A fast appearance based place recognition algorithm," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2017, pp. 5702–5708.
- [43] S. Lowry, N. Sünderhauf, P. Newman, J. J. Leonard, D. Cox, P. Corke, and M. J. Milford, "Visual place recognition: A survey," *IEEE Trans. Robot.*, vol. 32, no. 1, pp. 1–19, Feb. 2015.
- [44] Z. Chen, O. Lam, A. Jacobson, and M. Milford, "Convolutional neural network-based place recognition," in *Proc. Australas. Conf. Robot. Autom.*, 2014.
- [45] N. Sünderhauf, S. Shirazi, F. Dayoub, B. Upcroft, and M. Milford, "On the performance of ConvNet features for place recognition," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2015, pp. 4297–4304.
- [46] R. Arandjelovic, P. Gronat, A. Torii, T. Pajdla, and J. Sivic, "NetVLAD: CNN architecture for weakly supervised place recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 5297–5307.
- [47] Z. Chen, A. Jacobson, N. Sünderhauf, B. Upcroft, L. Liu, C. Shen, I. Reid, and M. Milford, "Deep learning features at scale for visual place recognition," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2017, pp. 3223–3230.
- [48] T. Naseer, W. Burgard, and C. Stachniss, "Robust visual localization across seasons," *IEEE Trans. Robot.*, vol. 34, no. 2, pp. 289–302, Apr. 2018.
- [49] S. Hausler, A. Jacobson, and M. Milford, "Multi-process fusion: Visual place recognition using multiple image processing methods," *IEEE Robot. Autom. Lett.*, vol. 4, no. 2, pp. 1924–1931, Feb. 2019.
- [50] K. Granstrom, J. Callmer, F. Ramos, and J. Nieto, "Learning to detect loop closure from range data," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2009, pp. 15–22.
- [51] M. Himstedt, J. Frost, S. Hellbach, H.-J. Böhme, and E. Maehle, "Large scale place recognition in 2D LIDAR scans using geometrical landmark relations," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sep. 2014, pp. 5030–5035.
- [52] B. Steder, G. Grisetti, and W. Burgard, "Robust place recognition for 3D range data based on point features," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2010, pp. 1400–1405.
- [53] R. Dubé, D. Dugas, E. Stumm, J. Nieto, R. Siegwart, and C. Cadena, "SegMatch: Segment based place recognition in 3D point clouds," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2017, pp. 5266–5272.
- [54] H. Yin, X. Ding, L. Tang, Y. Wang, and R. Xiong, "Efficient 3D LIDAR based loop closing using deep neural network," in *Proc. IEEE Int. Conf. Robot. Biomimetics (ROBIO)*, Dec. 2017, pp. 481–486.
- [55] R. C. Smith and P. Cheeseman, "On the representation and estimation of spatial uncertainty," *Int. J. Robot. Res.*, vol. 5, no. 4, pp. 56–68, 1986.
- [56] R. Smith, M. Self, and P. Cheeseman, "Estimating uncertain spatial relationships in robotics," in *Autonomous Robot Vehicles*, I. J. Cox and G. T. Wilfong, Eds. New York, NY, USA: Springer, 1990, pp. 167–193.
- [57] J. J. Leonard and H. F. Durrant-Whyte, "Simultaneous map building and localization for an autonomous mobile robot," in *Proc. IEEE/RSJ Int. Workshop Intell. Robots Syst. (IROS)*, Nov. 1991, pp. 1442–1447.
- [58] M. W. M. G. Dissanayake, P. Newman, S. Clark, H. F. Durrant-Whyte, and M. Csorba, "A solution to the simultaneous localization and map building (SLAM) problem," *IEEE Trans. Robot. Autom.*, vol. 17, no. 3, pp. 229–241, Jun. 2001.
- [59] S. Thrun, Y. Liu, D. Koller, A. Y. Ng, Z. Ghahramani, and H. Durrant-Whyte, "Simultaneous localization and mapping with sparse extended information filters," *Int. J. Robot. Res.*, vol. 23, nos. 7–8, pp. 693–716, 2004.
- [60] A. Eliazar and R. Parr, "DP-SLAM: Fast, robust simultaneous localization and mapping without predetermined landmarks," in *Proc. IJCAI, Acapulco, Mexico*, vol. 3, 2003, pp. 1135–1142.

- [61] M. Montemerlo, S. Thrun, D. Koller, and B. Wegbreit, "FastSLAM: A factored solution to the simultaneous localization and mapping problem," in *Proc. 18th AAAI Nat. Conf. Artif. Intell.* Edmonton, AB, USA: AAAI, 2002, pp. 593–598.
- [62] M. Montemerlo, S. Thrun, D. Roller, and B. Wegbreit, "FastSLAM 2.0: An improved particle filtering algorithm for simultaneous localization and mapping that provably converges," in *Proc. 18th Int. Joint Conf. Artif. Intell.*, 2003, pp. 1151–1156.
- [63] G. Grisetti, C. Stachniss, and W. Burgard, "Improved techniques for grid mapping with rao-blackwellized particle filters," *IEEE Trans. Robot.*, vol. 23, no. 1, pp. 34–46, Feb. 2007.
- [64] S. Thrun and M. Montemerlo, "The graph SLAM algorithm with applications to large-scale mapping of urban structures," *Int. J. Robot. Res.*, vol. 25, nos. 5–6, pp. 403–429, 2006.
- [65] R. Kümmerle, G. Grisetti, H. Strasdat, K. Konolige, and W. Burgard, "G²: A general framework for graph optimization," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2011, pp. 3607–3613.
- [66] F. Dellaert and M. Kaess, "Square root SAM: Simultaneous localization and mapping via square root information smoothing," *Int. J. Robot. Res.*, vol. 25, no. 12, pp. 1181–1203, 2006.
- [67] M. Kaess, A. Ranganathan, and F. Dellaert, "iSAM: Incremental smoothing and mapping," *IEEE Trans. Robot.*, vol. 24, no. 6, pp. 1365–1378, Dec. 2008.
- [68] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J. Leonard, and F. Dellaert, "iSAM2: Incremental smoothing and mapping using the Bayes tree," *Int. J. Robot. Res.*, vol. 31, no. 2, pp. 216–235, Feb. 2012.
- [69] J. A. Castellanos, J. Neira, and J. D. Tardós, "Limits to the consistency of EKF-based SLAM," *IFAC Proc. Volumes*, vol. 37, no. 8, pp. 716–721, Jul. 2004.
- [70] S. J. Julier and J. K. Uhlmann, "A counter example to the theory of simultaneous localization and map building," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2001, pp. 4238–4243.
- [71] S. Huang and G. Dissanayake, "Convergence and consistency analysis for extended Kalman filter based SLAM," *IEEE Trans. Robot.*, vol. 23, no. 5, pp. 1036–1049, Oct. 2007.
- [72] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, "State estimation in discrete-time linear dynamic systems," in *Estimation With Applications to Tracking and Navigation*. Hoboken, NJ, USA: Wiley, 2001, ch. 5, pp. 199–266.
- [73] S. Huang and G. Dissanayake, "A critique of current developments in simultaneous localization and mapping," *Int. J. Adv. Robotic Syst.*, vol. 13, no. 5, 2016, Art. no. 1729881416669482.
- [74] G. Grisetti, T. Guadagnino, I. Aloise, M. Colosi, B. Della Corte, and D. Schlegel, "Least squares optimization: From theory to practice," *Robotics*, vol. 9, no. 3, p. 51, Jul. 2020.
- [75] P. Agarwal, G. D. Tipaldi, L. Spinello, C. Stachniss, and W. Burgard, "Robust map optimization using dynamic covariance scaling," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2013, pp. 62–69.
- [76] N. Sunderhauf and P. Protzel, "Switchable constraints for robust pose graph SLAM," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2012, pp. 1879–1884.
- [77] N. Sunderhauf and P. Protzel, "Towards a robust back-end for pose graph SLAM," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2012, pp. 1254–1261.
- [78] E. Olson and P. Agarwal, "Inference on networks of mixtures for robust robot mapping," *Int. J. Robot. Res.*, vol. 32, no. 7, pp. 826–840, 2013.
- [79] N. Chebrou, T. Låbe, O. Vysotska, J. Behley, and C. Stachniss, "Adaptive robust kernels for non-linear least squares problems," *IEEE Robot. Autom. Lett.*, vol. 6, no. 2, pp. 2240–2247, Apr. 2021.
- [80] Y. Latif, C. Cadená, and J. Neira, "Robust loop closing over time for pose graph SLAM," *Int. J. Robot. Res.*, vol. 32, no. 14, pp. 1611–1626, Oct. 2013.
- [81] D. Scaramuzza and F. Fraundorfer, "Visual odometry [tutorial]," *IEEE Robot. Autom. Mag.*, vol. 18, no. 4, pp. 80–92, Dec. 2011.
- [82] T. Taketomi, H. Uchiyama, and S. Ikeda, "Visual SLAM algorithms: A survey from 2010 to 2016," *IPSA Trans. Comput. Vis. Appl.*, vol. 9, no. 1, p. 16, Dec. 2017.
- [83] A. J. Davison and D. W. Murray, "Simultaneous localization and map-building using active vision," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 865–880, Jul. 2002.
- [84] S. B. Goldberg, M. W. Maimone, and L. Matthies, "Stereo vision and rover navigation software for planetary exploration," in *Proc. IEEE Aerosp. Conf.*, vol. 5, Mar. 2002, p. 5.
- [85] A. J. Davison and N. Kita, "3D simultaneous localisation and map-building using active vision for a robot moving on undulating terrain," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 1, Dec. 2001, p. 1.
- [86] A. Howard, "Real-time stereo visual odometry for autonomous ground vehicles," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sep. 2008, pp. 3946–3952.
- [87] M. Sanfourche, V. Vittori, and G. L. Besnerais, "EVO: A realtime embedded stereo odometry for MAV applications," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Nov. 2013, pp. 2107–2114.
- [88] R. Mur-Artal and J. D. Tardós, "ORB-SLAM2: An open-source slam system for monocular, stereo, and RGB-D cameras," *IEEE Trans. Robot.*, vol. 33, no. 5, pp. 1255–1262, Oct. 2017.
- [89] A. J. Davison, "Real-time simultaneous localisation and mapping with a single camera," in *Proc. 9th IEEE Int. Conf. Comput. Vis.*, Oct. 2003, pp. 1403–1410.
- [90] A. J. Davison, I. D. Reid, N. D. Molton, and O. Stasse, "MonoSLAM: Real-time single camera SLAM," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 6, pp. 1052–1067, Jun. 2007.
- [91] J. Engel, J. Stückler, and D. Cremers, "Large-scale direct SLAM with stereo cameras," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2015, pp. 1935–1942.
- [92] J. Engel, V. Koltun, and D. Cremers, "Direct sparse odometry," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 3, pp. 611–625, Mar. 2018.
- [93] R. Mur-Artal, J. M. M. Montiel, and J. D. Tardós, "ORB-SLAM: A versatile and accurate monocular SLAM system," *IEEE Trans. Robot.*, vol. 31, no. 5, pp. 1147–1163, Oct. 2015.
- [94] H. Strasdat, J. M. M. Montiel, and A. J. Davison, "Real-time monocular SLAM: Why filter?" in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2010, pp. 2657–2664.
- [95] J. Engel, T. Schöps, and D. Cremers, "LSD-SLAM: Large-scale direct monocular SLAM," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 834–849.
- [96] K. Konda and R. Memisevic, "Learning visual odometry with a convolutional network," in *Proc. 10th Int. Conf. Comput. Vis. Theory Appl.*, 2015, pp. 486–490.
- [97] S. Wang, R. Clark, H. Wen, and N. Trigoni, "DeepVO: Towards end-to-end visual odometry with deep recurrent convolutional neural networks," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2017, pp. 2043–2050.
- [98] S. Wang, R. Clark, H. Wen, and N. Trigoni, "End-to-end, sequence-to-sequence probabilistic visual odometry through deep neural networks," *Int. J. Robot. Res.*, vol. 37, nos. 4–5, pp. 513–542, 2017.
- [99] Y. Almalioglu, M. R. U. Saputra, P. P. B. D. Gusmao, A. Markham, and N. Trigoni, "GANVO: Unsupervised deep monocular visual odometry and depth estimation with generative adversarial networks," in *Proc. Int. Conf. Robot. Autom. (ICRA)*, May 2019, pp. 5474–5480.
- [100] P. Muller and A. Savakis, "Flowdometry: An optical flow and deep learning based approach to visual odometry," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Mar. 2017, pp. 624–631.
- [101] G. Costante and T. A. Ciarfuglia, "LS-VO: Learning dense optical subspace for robust visual odometry estimation," *IEEE Robot. Autom. Lett.*, vol. 3, no. 3, pp. 1735–1742, Jul. 2018.
- [102] S. Edwards, L. Mihaylova, J. Aitken, and S. Anderson, "Toward robust visual odometry using prior 2D map information," in *Towards Autonomous Robotic Systems*. Cham, Switzerland: Springer, 2021.
- [103] B. Teixeira, H. Silva, A. Matos, and E. Silva, "Deep learning for underwater visual odometry estimation," *IEEE Access*, vol. 8, pp. 44687–44701, 2020.
- [104] C. Zhao, L. Sun, P. Purkait, T. Duckett, and R. Stolkin, "Learning monocular visual odometry with dense 3D mapping from dense 3D flow," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2018, pp. 6864–6871.
- [105] D. Strelow and S. Singh, "Motion estimation from image and inertial measurements," *Int. J. Robot. Res.*, vol. 23, no. 12, pp. 1157–1195, Dec. 2004.
- [106] G. Huang, M. Kaess, and J. J. Leonard, "Towards consistent visual-inertial navigation," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2014, pp. 4926–4933.
- [107] A. I. Mourikis and S. I. Roumeliotis, "A multi-state constraint Kalman filter for vision-aided inertial navigation," in *Proc. IEEE Int. Conf. Robot. Autom.*, Apr. 2007, pp. 3565–3572.
- [108] M. Bloesch, S. Omari, M. Hutter, and R. Siegwart, "Robust visual inertial odometry using a direct EKF-based approach," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2015, pp. 298–304.

- [109] T. Qin, P. Li, and S. Shen, "VINS-mono: A robust and versatile monocular visual-inertial state estimator," *IEEE Trans. Robot.*, vol. 34, no. 4, pp. 1004–1020, Aug. 2018.
- [110] L. Von Stumberg, V. Usenko, and D. Cremers, "Direct sparse visual-inertial odometry using dynamic marginalization," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 2510–2517.
- [111] J. Delmerico and D. Scaramuzza, "A benchmark comparison of monocular visual-inertial odometry algorithms for flying robots," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 2502–2509.
- [112] R. Clark, S. Wang, H. Wen, A. Markham, and N. Trigoni, "VINet: Visual-inertial odometry as a sequence-to-sequence learning problem," in *Proc. 31st AAAI Conf. Artif. Intell.* Palo Alto, CA, USA: AAAI Press, 2017, pp. 3995–4001.
- [113] F. Lu and E. Milios, "Globally consistent range scan alignment for environment mapping," *Auto. Robots*, vol. 4, no. 4, pp. 333–349, 1997.
- [114] A. Diosi and L. Kleeman, "Fast laser scan matching using polar coordinates," *Int. J. Robot. Res.*, vol. 26, no. 10, pp. 1125–1153, Oct. 2007.
- [115] W. Hess, D. Kohler, H. Rapp, and D. Andor, "Real-time loop closure in 2D LIDAR SLAM," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2016, pp. 1271–1278.
- [116] D. M. Cole and P. M. Newman, "Using laser range data for 3D SLAM in outdoor environments," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2006, pp. 1556–1563.
- [117] A. Nüchter, K. Lingemann, J. Hertzberg, and H. Surmann, "6D SLAM-3D mapping outdoor environments," *J. Field Robot.*, vol. 24, nos. 8–9, pp. 699–722, Aug. 2007.
- [118] D. Droschel and S. Behnke, "Efficient continuous-time SLAM for 3D lidar-based online mapping," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 5000–5007.
- [119] J. Zhang and S. Singh, "Low-drift and real-time lidar odometry and mapping," *Auton. Robots*, vol. 41, no. 2, pp. 401–416, Feb. 2017.
- [120] J. Zhang and S. Singh, "LOAM: Lidar odometry and mapping in real-time," in *Proc. Robot., Sci. Syst.*, vol. 2, no. 9. Berkeley, CA, USA: Univ. California, 2014.
- [121] F. Pomerleau, F. Colas, R. Siegwart, and S. Magnenat, "Comparing ICP variants on real-world data sets," *Auto. Robots*, vol. 34, no. 3, pp. 133–148, 2013.
- [122] V. Nguyen, S. Gächter, A. Martinelli, N. Tomatis, and R. Siegwart, "A comparison of line extraction algorithms using 2D range data for indoor mobile robotics," *Auto. Robots*, vol. 23, no. 2, pp. 97–111, 2007.
- [123] S. Kohlbrecher, O. von Stryk, J. Meyer, and U. Klingauf, "A flexible and scalable SLAM system with full 3D motion estimation," in *Proc. IEEE Int. Symp. Saf., Secur., Rescue Robot.*, Nov. 2011, pp. 155–160.
- [124] F. Martín, R. Triebel, L. Moreno, and R. Siegwart, "Two different tools for three-dimensional mapping: De-based scan matching and feature-based loop detection," *Robotica*, vol. 32, no. 1, pp. 19–41, Jan. 2014.
- [125] K. Granström, T. B. Schön, J. I. Nieto, and F. T. Ramos, "Learning to close loops from range data," *Int. J. Robot. Res.*, vol. 30, no. 14, pp. 1728–1754, Dec. 2011.
- [126] M. Bosse and R. Zlot, "Map matching and data association for large-scale two-dimensional laser scan-based slam," *Int. J. Robot. Res.*, vol. 27, no. 6, pp. 667–691, 2008.
- [127] J. Zhang and S. Singh, "Visual-lidar odometry and mapping: Low-drift, robust, and fast," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2015, pp. 2174–2181.
- [128] C. Debeunne and D. Vivet, "A review of visual-LiDAR fusion based simultaneous localization and mapping," *Sensors*, vol. 20, no. 7, p. 2068, 2020.
- [129] M. Hsiao and M. Kaess, "MH-iSAM2: Multi-hypothesis iSAM using Bayes tree and hypo-tree," in *Proc. Int. Conf. Robot. Autom. (ICRA)*, May 2019, pp. 1274–1280.
- [130] G. Mills, A. Jackson, and R. Richardson, "Advances in the inspection of unpiggable pipelines," *Robotics*, vol. 6, no. 4, p. 36, Nov. 2017.
- [131] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. Alvey Vis. Conf.*, 1988, pp. 147–151.
- [132] E. Rosten and T. Drummond, "Machine learning for high-speed corner detection," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2006, pp. 430–443.
- [133] M. Calonder, V. Lepetit, C. Strecha, and P. Fua, "BRIEF: Binary robust independent elementary features," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2010, pp. 778–792.
- [134] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 2564–2571.
- [135] J. J. Leonard and H. F. Durrant-Whyte, "Mobile robot localization by tracking geometric beacons," *IEEE Trans. Robot. Automat.*, vol. 7, no. 3, pp. 376–382, Jun. 1991.
- [136] P. Jensfelt, D. J. Austin, O. Wijk, and M. Andersson, "Feature based CONDENSATION for mobile robot localization," in *Proc. IEEE Int. Conf. Robot. Autom. Symposia (ICRA)*, Apr. 2000, pp. 2531–2537.
- [137] A. Garulli, A. Giannitrapani, A. Rossi, and A. Vicino, "Mobile robot SLAM for line-based environment representation," in *Proc. 44th IEEE Conf. Decis. Control*, Dec. 2005, pp. 2041–2046.
- [138] Y.-H. Choi, T.-K. Lee, and S.-Y. Oh, "A line feature based SLAM with low grade range sensors using geometric constraints and active exploration for mobile robot," *Auto. Robots*, vol. 24, no. 1, pp. 13–27, Jan. 2008.
- [139] P. Hansen, H. Alismail, P. Rander, and B. Browning, "Visual mapping for natural gas pipe inspection," *Int. J. Rob. Res.*, vol. 34, nos. 4–5, pp. 532–538, 2015.
- [140] L. Pedraza, G. Dissanayake, J. V. Miro, D. Rodriguez-Losada, and F. Matia, "BS-SLAM: Shaping the world," in *Proc. Robot., Sci. Syst.*, vol. 29, Jun. 2007, pp. 1–8.
- [141] L. Pedraza, D. Rodriguez-Losada, F. Matia, G. Dissanayake, and J. V. Miró, "Extending the limits of feature-based SLAM with B-splines," *IEEE Trans. Robot.*, vol. 25, no. 2, pp. 353–366, Apr. 2009.
- [142] R. Simmons and S. Koenig, "Probabilistic robot navigation in partially observable environments," in *Proc. Int. Joint Conf. Artif. Intell.*, vol. 95, 1995, pp. 1080–1087.
- [143] A. R. Cassandra, L. P. Kaelbling, and J. A. Kurien, "Acting under uncertainty: Discrete Bayesian models for mobile-robot navigation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, vol. 2, Nov. 1996, pp. 963–972.
- [144] D. Fox, W. Burgard, and S. Thrun, "Markov localization for mobile robots in dynamic environments," *J. Artif. Intell. Res.*, vol. 11, pp. 391–427, Nov. 1999.
- [145] A. Elfes, "Using occupancy grids for mobile robot perception and navigation," *IEEE Computer*, vol. 22, no. 6, pp. 46–57, Jun. 1989.
- [146] H. P. Moravec, "Sensor fusion in certainty grids for mobile robots," *Sensor Devices Syst. Robot.*, vol. 9, no. 2, pp. 253–276, 1989.
- [147] J. Ryde and H. Hu, "3D mapping with multi-resolution occupied voxel lists," *Auto. Robots*, vol. 28, no. 2, p. 169, 2010.
- [148] W. Burgard, A. Derr, D. Fox, and A. B. Cremers, "Integrating global position estimation and position tracking for mobile robots: The dynamic Markov localization approach," in *Proc. IEEE Int. Conf. Intell. Robots Syst.*, vol. 2, Oct. 1998, pp. 730–735.
- [149] P. Payeur, P. Hébert, D. Laurendeau, and C. M. Gosselin, "Probabilistic octree modeling of a 3D dynamic environment," in *Proc. Int. Conf. Robot. Autom.*, vol. 2, Apr. 1997, pp. 1289–1296.
- [150] A. Hornung, K. M. Wurm, M. Bennewitz, C. Stachniss, and W. Burgard, "OctoMap: An efficient probabilistic 3D mapping framework based on octrees," *Auton. Robot.*, vol. 34, no. 3, pp. 189–206, 2013.
- [151] D. Borrmann, J. Elseberg, K. Lingemann, A. Nüchter, and J. Hertzberg, "Globally consistent 3D mapping with scan matching," *Robot. Auto. Syst.*, vol. 56, no. 2, pp. 130–142, 2008.
- [152] H. Pfister, M. Zwicker, J. van Baar, and M. Gross, "Surfels: Surface elements as rendering primitives," in *Proc. 27th Annu. Conf. Comput. Graph. Interact. Techn. (SIGGRAPH)*, 2000, pp. 335–342.
- [153] T. Whelan, S. Leutenegger, R. F. Salas-Moreno, B. Glocker, and A. J. Davison, "ElasticFusion: Dense SLAM without a pose graph," in *Proc. Robot., Sci. Syst.*, 2015, pp. 1–9.
- [154] J. Behley and C. Stachniss, "Efficient surfel-based SLAM using 3D laser range data in urban environments," *Robot., Sci. Syst.*, vol. 2018, Jun. 2018.
- [155] R. A. Newcombe, A. Fitzgibbon, S. Izadi, O. Hilliges, D. Molyneaux, D. Kim, A. J. Davison, P. Kohi, J. Shotton, and S. Hodges, "KinectFusion: Real-time dense surface mapping and tracking," in *Proc. 10th IEEE Int. Symp. Mixed Augmented Reality*, Oct. 2011, pp. 127–136.
- [156] J. Boal, Á. Sánchez-Mirallas, and Á. Arranz, "Topological simultaneous localization and mapping: A survey," *Robotica*, vol. 32, no. 5, pp. 803–821, Aug. 2014.
- [157] D. Kortenkamp and T. Weymouth, "Topological mapping for mobile robots using a combination of sonar and vision sensing," in *Proc. Nat. Conf. Artif. Intell.*, vol. 2, 1994, pp. 979–984.
- [158] H. Choset and K. Nagatani, "Topological simultaneous localization and mapping (SLAM): Toward exact localization without explicit localization," *IEEE Trans. Robot. Autom.*, vol. 17, no. 2, pp. 125–137, Apr. 2001.

- [159] I. Nourbakhsh, R. Powers, and S. Birchfield, "DERVISH an office-navigating robot," *AI Mag.*, vol. 16, no. 2, pp. 53–60, 1995.
- [160] B. Kuipers and Y.-T. Byun, "A robot exploration and mapping strategy based on a semantic hierarchy of spatial representations," *Robot. Auto. Syst.*, vol. 8, nos. 1–2, pp. 47–63, Nov. 1991.
- [161] R. Worley and S. Anderson, "Topological robot localization in a large-scale water pipe network," in *Towards Autonomous Robotic Systems*, A. Mohammad, X. Dong, and M. Russo, Eds. Cham, Switzerland: Springer, 2020, pp. 77–89.
- [162] R. Worley and S. Anderson, "Robust efficient localization of robots in pipe networks using a particle filter for hybrid metric-topological space," in *Proc. Eur. Conf. Mobile Robots (ECMR)*, 2021.
- [163] A. Nüchter and J. Hertzberg, "Towards semantic maps for mobile robots," *Robot. Auto. Syst.*, vol. 56, no. 11, pp. 915–926, 2008.
- [164] I. Kostavelis and A. Gasteratos, "Semantic mapping for mobile robotics tasks: A survey," *Robot. Auton. Syst.*, vol. 66, pp. 86–103, Apr. 2015.
- [165] E. Rome, H. Surmann, H. Streich, U. Licht, and K.-L. Paap, "A custom IR scanner for landmark detection with the autonomous sewer robot MAKRO," in *Proc. Int. Symp. Intell. Robotic Syst.*, 2001, pp. 457–466.
- [166] D.-H. Lee, H. Moon, and H. R. Choi, "Landmark detection of in-pipe working robot using line-laser beam projection," in *Proc. ICCAS*, Oct. 2010, pp. 611–615.
- [167] S. H. Kim, S. J. Lee, and S. W. Kim, "Weaving laser vision system for navigation of mobile robots in pipeline structures," *IEEE Sensors J.*, vol. 18, no. 6, pp. 2585–2591, Mar. 2018.
- [168] J.-S. Lee, S.-G. Roh, D. W. Kim, H. Moon, and H. R. Choi, "In-pipe robot navigation based on the landmark recognition system using shadow images," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2009, pp. 1857–1862.
- [169] Y. S. Choi, H. M. Kim, H. M. Mun, Y. G. Lee, and H. R. Choi, "Recognition of pipeline geometry by using monocular camera and PSD sensors," *Intell. Service Robot.*, vol. 10, no. 3, pp. 213–227, Jul. 2017.
- [170] J.-L. Blanco, J.-A. Fernández-Madrigrá, and J. Gonzalez, "A new approach for large-scale localization and mapping: Hybrid metric-topological SLAM," in *Proc. IEEE Int. Conf. Robot. Autom.*, Apr. 2007, pp. 2061–2067.
- [171] K. Konolige, E. Marder-Eppstein, and B. Marthi, "Navigation in hybrid metric-topological maps," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2011, pp. 3041–3047.
- [172] P. Schmuck, S. A. Scherer, and A. Zell, "Hybrid metric-topological 3D occupancy grid maps for large-scale mapping," *IFAC-PapersOnLine*, vol. 49, no. 15, pp. 230–235, 2016.
- [173] R. F. Salas-Moreno, R. A. Newcombe, H. Strasdat, P. H. J. Kelly, and A. J. Davison, "SLAM++: Simultaneous localisation and mapping at the level of objects," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 1352–1359.
- [174] S. L. Bowman, N. Atanasov, K. Daniilidis, and G. J. Pappas, "Probabilistic data association for semantic SLAM," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2017, pp. 1722–1729.
- [175] B. Kuipers, J. Modayil, P. Beeson, M. Macmahon, and F. Savelli, "Local metrical and global topological maps in the hybrid spatial semantic hierarchy," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, Apr. 2004, pp. 4845–4851.
- [176] D.-H. Lee, H. Moon, J. C. Koo, and H. R. Choi, "Map building method for urban gas pipelines based on landmark detection," *Int. J. Control. Autom. Syst.*, vol. 11, no. 1, pp. 127–135, Feb. 2013.
- [177] P. L. Hanna, M. E. Napier, and V. Ashkenazi, "Strapdown inertial surveying for internal pipeline surveys," in *Kinematic Systems in Geodesy, Surveying, and Remote Sensing*. New York, NY, USA: Springer, 1991, pp. 140–153.
- [178] E.-H. Shin and N. El-Sheimy, "Navigation Kalman filter design for pipeline pigging," *J. Navigat.*, vol. 58, no. 2, pp. 283–295, 2005.
- [179] A. C. Murtra and J. M. M. Tur, "IMU and cable encoder data fusion for in-pipe mobile robot localization," in *Proc. IEEE Conf. Technol. Practical Robot Appl. (TePRA)*, Apr. 2013, pp. 1–6.
- [180] W. M. F. Al-Masri, M. F. Abdel-Hafez, and M. A. Jaradat, "Inertial navigation system of pipeline inspection gauge," *IEEE Trans. Control Syst. Technol.*, vol. 28, no. 2, pp. 609–616, Mar. 2020.
- [181] Q. Chen, Q. Zhang, X. Niu, and Y. Wang, "Positioning accuracy of a pipeline surveying system based on MEMS IMU and odometer: Case study," *IEEE Access*, vol. 7, pp. 104453–104461, 2019.
- [182] Y. Yang, L. Bin, W. Xinjie, and Y. Lijian, "Application of adaptive cubature Kalman filter to in-pipe survey system for 3D small-diameter pipeline mapping," *IEEE Sensors J.*, vol. 20, no. 12, pp. 6331–6337, Jun. 2020.
- [183] H. Sahli and N. El-Sheimy, "A novel method to enhance pipeline trajectory determination using pipeline junctions," *Sensors*, vol. 16, no. 4, pp. 1–17, 2016.
- [184] L. Guan, Y. Gao, A. Noureldin, and X. Cong, "Junction detection based on CCWT and MEMS accelerometer measurement," in *Proc. IEEE Int. Conf. Mechatronics Autom. (ICMA)*, Aug. 2018, pp. 1227–1232.
- [185] Y. Wu, E. Mittmann, C. Winston, and K. Youcef-Toumi, "A practical minimalism approach to in-pipe robot localization," in *Proc. Amer. Control Conf. (ACC)*, Jul. 2019, pp. 3180–3187.
- [186] D. Wu, D. Chatzigeorgiou, K. Youcef-Toumi, and R. Ben-Mansour, "Node localization in robotic sensor networks for pipeline inspection," *IEEE Trans. Ind. Informat.*, vol. 12, no. 2, pp. 809–819, Apr. 2015.
- [187] M. S. Chowdhury and M. F. Abdel-Hafez, "Pipeline inspection gauge position estimation using inertial measurement unit, odometer, and a set of reference stations," *ASCE-ASME J. Risk Uncert. Engng Syst. B, Mech. Engrg.*, vol. 2, no. 2, p. 21001, Jun. 2016.
- [188] E. H. A. Duisterwinkel, E. Talmishnikh, D. Krijnders, and H. J. Wortche, "Sensor motes for the exploration and monitoring of operational pipelines," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 3, pp. 655–666, Mar. 2018.
- [189] E. Rome, J. Hertzberg, F. Kirchner, U. Licht, and T. Christaller, "Towards autonomous sewer robots: The MAKRO project," *Urban Water*, vol. 1, no. 1, pp. 57–70, Mar. 1999.
- [190] A. A. F. Nassiraei, Y. Kawamura, A. Ahrary, Y. Mikuriya, and K. Ishii, "A new approach to the sewer pipe inspection: Fully autonomous mobile robot 'KANTARO,'" in *Proc. 32nd Annu. Conf. IEEE Ind. Electron. (IECON)*, Nov. 2006, pp. 4088–4093.
- [191] S.-G. Roh, D. W. Kim, J.-S. Lee, H. Moon, and H. R. Choi, "Modularized in-pipe robot capable of selective navigation inside of pipelines," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sep. 2008, pp. 1724–1729.
- [192] P. Debenest, M. Guarnieri, and S. Hirose, "PipeTron series—Robots for pipe inspection," in *Proc. 3rd Int. Conf. Appl. Robot. Power Ind.*, Oct. 2014, pp. 1–6.
- [193] H. Schempf, E. Mutschler, A. Gavaert, G. Skoptsov, and W. Crowley, "Visual and nondestructive evaluation inspection of live gas mains using the explorer family of pipe robots," *J. Field Robot.*, vol. 27, no. 3, pp. 217–249, May 2010.
- [194] D. Krysz and H. Najjaran, "Development of visual simultaneous localization and mapping (VSLAM) for a pipe inspection robot," in *Proc. Int. Symp. Comput. Intell. Robot. Autom.*, Jun. 2007, pp. 344–349.
- [195] P. Hansen, H. Alismail, B. Browning, and P. Rander, "Stereo visual odometry for pipe mapping," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sep. 2011, pp. 4020–4025.
- [196] P. Hansen, H. Alismail, P. Rander, and B. Browning, "Pipe mapping with monocular fisheye imagery," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Nov. 2013, pp. 5180–5185.
- [197] S. Esquivel, R. Koch, and H. Rehse, "Reconstruction of sewer shaft profiles from fisheye-lens camera images," in *Proc. Joint Pattern Recognit. Symp.*, 2009, pp. 332–341.
- [198] Y. Zhang, R. Hartley, J. Mashford, L. Wang, and S. Burn, "Pipeline reconstruction from fisheye images," *J. WSCG*, vol. 19, nos. 1–3, pp. 49–57, 2011.
- [199] J. Kunzel, T. Werner, P. Eisert, and J. Waschnowski, "Automatic analysis of sewer pipes based on unrolled monocular fisheye images," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Mar. 2018, pp. 2019–2027.
- [200] X. Zhang, P. Zhao, Q. Hu, H. Wang, M. Ai, and J. Li, "A 3D reconstruction pipeline of urban drainage pipes based on MultiviewImage matching using low-cost panoramic video cameras," *Water*, vol. 11, no. 10, p. 2101, Oct. 2019.
- [201] Y. Gong, R. S. Johnston, C. D. Melville, and E. J. Seibel, "Axial-stereo 3-D optical metrology for inner profile of pipes using a scanning laser endoscope," *Int. J. Optomechatronics*, vol. 9, no. 3, pp. 238–247, Jul. 2015, doi: 10.1080/15599612.2015.1059535.
- [202] P. Hansen, H. Alismail, P. Rander, and B. Browning, "Monocular visual odometry for robot localization in LNG pipes," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2011, pp. 3111–3116.
- [203] D. Alejo, F. Caballero, and L. Merino, "RGBD-based robot localization in sewer networks," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2017, pp. 4070–4076.

- [204] D. Alejo, F. Caballero, and L. Merino, "A robust localization system for inspection robots in sewer networks," *Sensors*, vol. 19, no. 22, p. 4946, Nov. 2019.
- [205] T.-Y. Chuang and C.-C. Sung, "Learning and SLAM based decision support platform for sewer inspection," *Remote Sens.*, vol. 12, no. 6, p. 968, Mar. 2020.
- [206] R. Zhang, M. H. Evans, R. Worley, S. R. Anderson, and L. Mihaylova, "Improving SLAM in pipe networks by leveraging cylindrical regularity," in *Towards Autonomous Robotic Systems*. Cham, Switzerland: Springer, 2021.
- [207] S. Kagami, H. Taira, N. Miyashita, A. Torii, and M. Okutomi, "3D pipe network reconstruction based on structure from motion with incremental conic shape detection and cylindrical constraint," in *Proc. IEEE 29th Int. Symp. Ind. Electron. (ISIE)*, Jun. 2020, pp. 1345–1352.
- [208] I. J. H. G. Vreeburg and D. J. B. Boxall, "Discolouration in potable water distribution systems: A review," *Water Res.*, vol. 41, no. 3, pp. 519–529, Feb. 2007.
- [209] S. E. Kahi, D. Asmar, A. Fakh, J. Nieto, and E. Nebot, "A vision-based system for mapping the inside of a pipe," in *Proc. IEEE Int. Conf. Robot. Biomimetics*, no. 1, Dec. 2011, pp. 2605–2611.
- [210] J. T. Thielemann, G. M. Breivik, and A. Berge, "Pipeline landmark detection for autonomous robot navigation using time-of-flight imagery," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2008, pp. 1–7.
- [211] J. Hertzberg and F. Kirchner, "Landmark-based autonomous navigation in sewerage pipes," in *Proc. 1st Euromicro Workshop Adv. Mobile Robots (EUROBOT)*, 1996, pp. 68–73.
- [212] O. Adria, H. Streich, and J. Hertzberg, "Dynamic replanning in uncertain environments for a sewer inspection robot," *Int. J. Adv. Robotic Syst.*, vol. 1, no. 1, p. 4, Mar. 2004.
- [213] Z. Liu and D. Kryz, "The use of laser range finder on a robotic platform for pipe inspection," *Mech. Syst. Signal Process.*, vol. 31, pp. 246–257, Aug. 2012.
- [214] A. Ahrary, Y. Kawamura, and M. Ishikawa, "A laser scanner for landmark detection with the sewer inspection robot KANTARO," in *Proc. IEEE/SMC Int. Conf. Syst. Syst. Eng.*, Apr. 2006, pp. 291–296.
- [215] D.-H. Lee, H. Moon, and H. R. Choi, "Autonomous navigation of in-pipe working robot in unknown pipeline environment," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2011, pp. 1559–1564.
- [216] D.-H. Lee, H. Moon, and H. R. Choi, "Landmark detection methods for in-pipe robot traveling in urban gas pipelines," *Robotica*, vol. 34, no. 3, pp. 601–618, Mar. 2016.
- [217] A. Ahrary and M. Ishikawa, "Self-localization of autonomous sewer robots by using a stereo camera and a laser scanner," in *Proc. IEEE Int. Conf. Netw., Sens. Control*, Apr. 2006, pp. 78–83.
- [218] K. Ma, J. Zhu, T. Dodd, R. Collins, and S. Anderson, "Robot mapping and localisation for feature sparse water pipes using voids as landmarks," in *Towards Autonomous Robotic Systems (TAROS)* (Lecture Notes in Computer Science), vol. 9287. Cham, Switzerland: Springer, 2015, pp. 161–166.
- [219] K. Ma, M. Schirru, A. H. Zahraee, R. Dwyer-Joyce, J. Boxall, T. J. Dodd, R. Collins, and S. R. Anderson, "PipeSLAM: Simultaneous localisation and mapping in feature sparse water pipes using the rao-blackwellised particle filter," in *Proc. IEEE Int. Conf. Adv. Intell. Mechatronics (AIM)*, Jul. 2017, pp. 1459–1464.
- [220] K. Ma, M. M. Schirru, A. H. Zahraee, R. Dwyer-Joyce, J. Boxall, T. J. Dodd, R. Collins, and S. R. Anderson, "Robot mapping and localisation in metal water pipes using hydrophone induced vibration and map alignment by dynamic time warping," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2017, pp. 2548–2553.
- [221] R. Worley, K. Ma, G. Sailor, M. M. Schirru, R. Dwyer-Joyce, J. Boxall, T. Dodd, R. Collins, and S. Anderson, "Robot localization in water pipes using acoustic signals and pose graph optimization," *Sensors*, vol. 20, no. 19, p. 5584, Sep. 2020.
- [222] T. Seco, C. Rizzo, J. Espelósín, and J. L. Villarreal, "A robot localization system based on RF fading using particle filters inside pipes," in *Proc. Int. Conf. Auto. Robot Syst. Competitions (ICARSC)*, May 2016, pp. 28–34.
- [223] C. Rizzo, T. Seco, J. Espelósín, F. Lera, and J. L. Villarreal, "An alternative approach for robot localization inside pipes using RF spatial fading," *Robot. Auto. Syst.*, vol. 136, Feb. 2021, Art. no. 103702.
- [224] Y. Bando, H. Suhara, M. Tanaka, T. Kamegawa, K. Itoyama, K. Yoshii, F. Matsuno, and H. G. Okuno, "Sound-based Online localization for an in-pipe snake robot," in *Proc. IEEE Int. Symp. Saf., Secur., Rescue Robot. (SSRR)*, Oct. 2016, pp. 207–213.
- [225] R. Worley, Y. Yu, and S. Anderson, "Acoustic echo-localization for pipe inspection robots," in *Proc. IEEE Int. Conf. Multisensor Fusion Integr. Intell. Syst. (MFI)*, Sep. 2020, pp. 2–7.
- [226] L. P. Peters, J. J. Daniels, and J. D. Young, "Ground penetrating radar as a subsurface environmental sensing tool," *Proc. IEEE*, vol. 82, no. 12, pp. 1802–1822, Dec. 1994.
- [227] W. W.-L. Lai, X. Dérobert, and P. Annan, "A review of ground penetrating radar application in civil engineering: A 30-year journey from locating and testing to imaging and diagnosis," *NDT&E Int.*, vol. 96, pp. 58–78, Jun. 2018.
- [228] S. Li, H. Cai, D. M. Abraham, and P. Mao, "Estimating features of underground utilities: Hybrid GPR/GPS approach," *J. Comput. Civil Eng.*, vol. 30, no. 1, Jan. 2016, Art. no. 04014108.
- [229] X. Zhou, H. Chen, and T. Hao, "Efficient detection of buried plastic pipes by combining GPR and electric field methods," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 6, pp. 3967–3979, Jun. 2019.
- [230] H. Li, C. Chou, L. Fan, B. Li, D. Wang, and D. Song, "Robotic subsurface pipeline mapping with a ground-penetrating radar and a camera," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2018, pp. 3145–3150.
- [231] H. Li, C. Chou, L. Fan, B. Li, D. Wang, and D. Song, "Toward automatic subsurface pipeline mapping by fusing a ground-penetrating radar and a camera," *IEEE Trans. Autom. Sci. Eng.*, vol. 17, no. 2, pp. 722–734, Apr. 2019.
- [232] C. Chou, H. Li, and D. Song, "Encoder-camera-ground penetrating radar sensor fusion: Bimodal calibration and subsurface mapping," *IEEE Trans. Robot.*, vol. 37, no. 1, pp. 67–81, Feb. 2021.
- [233] Q. Dou, L. Wei, D. R. Magee, and A. Cohn, "Real-time hyperbola recognition and fitting in GPR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 1, pp. 51–62, Jan. 2016.
- [234] H. Chen and A. G. Cohn, "Buried utility pipeline mapping based on multiple spatial data sources: A Bayesian data fusion approach," in *Proc. 22nd Int. Joint Conf. Artif. Intell. (IJCAI)*, 2011, pp. 2411–2417.
- [235] M. Bilal, W. Khan, J. Muggleton, E. Rustighi, H. Jenks, S. R. Pennock, P. R. Atkins, and A. Cohn, "Inferring the most probable maps of underground utilities using Bayesian mapping model," *J. Appl. Geophys.*, vol. 150, pp. 52–66, Mar. 2018.
- [236] J. Cai, J. Jeon, H. Cai, and S. Li, "Fusing heterogeneous information for underground utility map generation based on Dempster-Shafer theory," *J. Comput. Civil Eng.*, vol. 34, no. 3, May 2020, Art. no. 04020013.
- [237] Q. Dou, L. Wei, D. R. Magee, P. R. Atkins, D. N. Chapman, G. Curioni, K. F. Goddard, F. Hayati, H. Jenks, N. Metje, and J. Muggleton, "3D buried utility location using a marching-cross-section algorithm for multi-sensor data fusion," *Sensors*, vol. 16, no. 11, p. 1827, Aug. 2016.
- [238] D. Alejo, F. Chataigner, D. Serrano, L. Merino, and F. Caballero, "Into the dirt: Datasets of sewer networks with aerial and ground platforms," *J. Field Robot.*, vol. 38, no. 1, pp. 105–120, Jan. 2021.
- [239] N. Metje, D. N. Chapman, D. Cheneler, M. Ward, and A. M. Thomas, "Smart pipes—Instrumented water pipes. can this be made a reality?" *Sensors*, vol. 11, no. 8, pp. 7455–7475, Jul. 2011.
- [240] G. Jia, G. Han, H. Rao, and L. Shu, "Edge computing-based intelligent manhole cover management system for smart cities," *IEEE Internet Things J.*, vol. 5, no. 3, pp. 1648–1656, Jun. 2018.
- [241] J.-H. Kim, G. Sharma, N. Boudriga, S. S. Iyengar, and N. Prabakar, "Autonomous pipeline monitoring and maintenance system: A RFID-based approach," *EURASIP J. Wireless Commun. Netw.*, vol. 2015, no. 1, p. 262, Dec. 2015.
- [242] W. Ji, M. Chen, J. Davies, Z. Xu, B. Wang, and K. Sun, "The development of a RFID-based smart pipeline tracking and information management system," in *Proc. 24th Int. Conf. Autom. Comput. (ICAC)*, Sep. 2018, pp. 1–6.
- [243] S. K. Devitt, *Trustworthiness of Autonomous Systems*. Cham, Switzerland: Springer, 2018, pp. 161–184, doi: [10.1007/978-3-319-64816-3_9](https://doi.org/10.1007/978-3-319-64816-3_9).
- [244] S. Shahrddar, L. Menezes, and M. Nojournian, "A survey on trust in autonomous systems," in *Proc. Comput. Conf.*, vol. 2, London, U.K., 2018, pp. 368–386.
- [245] M. Fisher, V. Mascardi, K. Yvonne Rozier, B.-H. Schlingloff, M. Winikoff, and N. Yorke-Smith, "Towards a framework for certification of reliable autonomous systems," 2020, *arXiv:2001.09124*. [Online]. Available: <http://arxiv.org/abs/2001.09124>

- [246] H. He, J. Gray, A. Cangelosi, Q. Meng, T. M. McGinnity, and J. Mehnen, "The challenges and opportunities of artificial intelligence for trust-worthy robots and autonomous systems," in *Proc. 3rd Int. Conf. Intell. Robotic Control Eng. (IRCE)*, Oxford, U.K., Aug. 2020, pp. 68–74.
- [247] E. Toreini, M. Aitken, K. Coopamootoo, K. Elliott, C. G. Zelaya, and A. van Moorsel, "The relationship between trust in AI and trustworthy machine learning technologies," in *Proc. Conf. Fairness, Accountability, Transparency*, Jan. 2020.
- [248] D. Cameron, S. de Saille, E. C. Collins, J. M. Aitken, H. Cheung, A. Chua, E. J. Loh, and J. Law, "The effect of social-cognitive recovery strategies on likability, capability and trust in social robots," *Comput. Hum. Behav.*, vol. 114, Jan. 2021, Art. no. 106561.
- [249] C. Parrott, T. J. Dodd, J. Boxall, and K. Horoshenkov, "Simulation of the behavior of biologically-inspired swarm robots for the autonomous inspection of buried pipes," *Tunnelling Underground Space Technol.*, vol. 101, Jul. 2020, Art. no. 103356.
- [250] C. Leung, S. Huang, and G. Dissanayake, "Active SLAM using model predictive control and attractor based exploration," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2006, pp. 5026–5031.
- [251] L. Carlone, J. Du, M. K. Ng, B. Bona, and M. Indri, "Active SLAM and exploration with particle filters using kullback-leibler divergence," *J. Intell. Robotic Syst.*, vol. 75, no. 2, pp. 291–311, Aug. 2014.
- [252] H. Carrillo, I. Reid, and J. A. Castellanos, "On the comparison of uncertainty criteria for active SLAM," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2012, pp. 2080–2087.
- [253] R. Valencia, M. Morta, J. Andrade-Cetto, and J. M. Porta, "Planning reliable paths with pose SLAM," *IEEE Trans. Robot.*, vol. 29, no. 4, pp. 1050–1059, Aug. 2013.
- [254] V. A. Tsihrintzis, R. Hamid, and H. R. Fuentes, "Use of geographic information systems (GIS) in water resources: A review," *Water Resour. Manage.*, vol. 10, no. 4, pp. 251–277, Aug. 1996.
- [255] U. M. Shamsi, *GIS Tools for Water, Wastewater, and Stormwater Systems*. Reston, VA, USA: ASCE Press, 2002.
- [256] I. Seth, "GIS: A useful tool in urban water management," in *Geospatial Tools for Urban Water Resources*, P. Lawrence, Ed. Dordrecht, The Netherlands: Springer, 2013.
- [257] T. Nan, M. Xianyang, L. Lin, C. Liu, and J. Xiuhan, "Manhole detection and location for urban pavement," in *Proc. IEEE Int. Conf. Intell. Transp. Syst.*, vol. 2, Oct. 2003, pp. 1552–1555.
- [258] R. Timofte and L. Van Gool, "Multi-view manhole detection, recognition, and 3D localisation," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops (ICCV Workshops)*, Nov. 2011, pp. 188–195.
- [259] Y. Yu, J. Li, H. Guan, C. Wang, and J. Yu, "Automated detection of road manhole and sewer well covers from mobile LiDAR point clouds," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 9, pp. 1549–1553, Sep. 2014.
- [260] Y. Yu, H. Guan, and Z. Ji, "Automated detection of urban road manhole covers using mobile laser scanning data," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 6, pp. 3258–3269, Dec. 2015.
- [261] Z. Wei, M. Yang, L. Wang, H. Ma, X. Chen, and R. Zhong, "Customized mobile LiDAR system for manhole cover detection and identification," *Sensors*, vol. 19, no. 10, p. 2422, May 2019.
- [262] M. Vitry, K. Schindler, J. Rieckermann, and J. Leitão, "Sewer inlet localization in UAV image clouds: Improving performance with multiview detection," *Remote Sens.*, vol. 10, no. 5, p. 706, May 2018.
- [263] B. Commandre, D. En-Nejjary, L. Pibre, M. Chaumont, C. Delenne, and N. Chahinian, "Manhole cover localization in aerial images with a deep learning approach," in *Proc. Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci. Hannover Workshop*, vol. 42, 2017, pp. 333–338.
- [264] W. Liu, D. Cheng, P. Yin, M. Yang, E. Li, M. Xie, and L. Zhang, "Small manhole cover detection in remote sensing imagery with deep convolutional neural networks," *ISPRS Int. J. Geo-Inf.*, vol. 8, no. 1, p. 49, Jan. 2019.
- [265] V. Krylov, E. Kenny, and R. Dahyot, "Automatic discovery and geotagging of objects from street view imagery," *Remote Sens.*, vol. 10, no. 5, p. 661, Apr. 2018.
- [266] B. Commandre, N. Chahinian, J.-S. Bailly, M. Chaumont, G. Subsol, F. Rodriguez, M. Derrass, L. Deruelle, and C. Delenne, "Automatic reconstruction of urban wastewater and stormwater networks based on uncertain manhole cover locations," in *Proc. Int. Conf. Urban Drainage*, 2017, pp. 2345–2352.
- [267] N. Chahinian, C. Delenne, B. Commandre, M. Derrass, L. Deruelle, and J.-S. Bailly, "Automatic mapping of urban wastewater networks based on manhole cover locations," *Comput., Environ. Urban Syst.*, vol. 78, Nov. 2019, Art. no. 101370.
- [268] S. Ardeshir, A. R. Zamir, A. Torroella, and M. Shah, "GIS-assisted object detection and geospatial localization," in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2014, pp. 602–617.
- [269] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part-based models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 9, pp. 1627–1645, Sep. 2009.
- [270] E. Waltz and T. Waltz, "Principles and practice of image and spatial data fusion," in *Handbook of Multisensor Data Fusion: Theory and Practice*, 2nd ed., M. Liggins, D. Hall, and J. Llinas, Eds. Boca Raton, FL, USA: CRC Press, 2017.
- [271] M. Kampffmeyer, A.-B. Salberg, and R. Jenssen, "Semantic segmentation of small objects and modeling of uncertainty in urban remote sensing images using deep convolutional neural networks," in *Proc. IEEE CVPR Workshops*, Jun. 2016, pp. 680–688.
- [272] C. Georgiou, S. Anderson, and T. Dodd, "Constructing informative Bayesian map priors: A multi-objective optimisation approach applied to indoor occupancy grid mapping," *Int. J. Robot. Res.*, vol. 36, no. 3, pp. 274–291, Mar. 2017.
- [273] S. Saeedi, M. Trentini, M. Set, and H. Li, "Multiple-robot simultaneous localization and mapping: A review," *J. Field Robot.*, vol. 7, pp. 81–86, Jan. 2015.
- [274] F. Castanedo, "A review of data fusion techniques," *Sci. World J.*, vol. 2013, pp. 1–19, Oct. 2013.
- [275] M. A. Bakr and S. Lee, "Distributed multisensor data fusion under unknown correlation and data inconsistency," *Sensors*, vol. 17, no. 11, p. 2472, Oct. 2017.
- [276] X. Zhou and S. Roumeliotis, "Multi-robot SLAM with unknown initial correspondence: The robot rendezvous case," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2006, pp. 1785–1792.
- [277] E. W. Nettleton, P. W. Gibbens, and H. F. Durrant-Whyte, "Closed form solutions to the multiple-platform simultaneous localization and map building (SLAM) problem," in *Proc. SPIE, 4th Sensor Fusion, Architectures, Algorithms, Appl.*, vol. 4051. Bellingham, WA, USA: SPIE, Apr. 2000, pp. 428–437.
- [278] S. Thrun and Y. Liu, "Multi-robot SLAM with sparse extended information filters," in *Proc. 11th Int. Symp. Robot. Res.*, 2005, pp. 254–266.
- [279] A. Howard, "Multi-robot simultaneous localization and mapping using particle filters," *Int. J. Robot. Res.*, vol. 25, no. 12, pp. 1243–1256, Dec. 2006.
- [280] B. Kim, M. Kaess, L. Fletcher, J. Leonard, A. Bachrach, N. Roy, and S. Teller, "Multiple relative pose graphs for robust cooperative mapping," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2010, pp. 3185–3192.
- [281] A. Cunningham, K. M. Wurm, W. Burgard, and F. Dellaert, "Fully distributed scalable smoothing and mapping with robust multi-robot data association," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2012, pp. 1093–1100.
- [282] A. Cunningham, V. Indelman, and F. Dellaert, "DDF-SAM 2.0: Consistent distributed smoothing and mapping," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2013, pp. 5220–5227.
- [283] R. Natarajan and M. A. Gennert, "Efficient factor graph fusion for multi-robot mapping and beyond," in *Proc. 21st Int. Conf. Inf. Fusion (FUSION)*, Jul. 2018, pp. 1137–1145.
- [284] C. Forster, S. Lynen, L. Kneip, and D. Scaramuzza, "Collaborative monocular SLAM with multiple micro aerial vehicles," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Nov. 2013, pp. 3962–3970.
- [285] L. Riazuelo, J. Civera, and J. M. M. Montiel, "C2TAM: A cloud framework for cooperative tracking and mapping," *Robot. Auto. Syst.*, vol. 62, no. 4, pp. 401–413, Apr. 2014.
- [286] H. Zhang, X. Chen, H. Lu, and J. Xiao, "Distributed and collaborative monocular simultaneous localization and mapping for multi-robot systems in large-scale environments," *Int. J. Adv. Robotic Syst.*, vol. 15, no. 3, pp. 1–20, 2018.
- [287] T. Cieslewski, S. Choudhary, and D. Scaramuzza, "Data-efficient decentralized visual SLAM," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 2466–2473.
- [288] J. G. Mangelson, D. Dominic, R. M. Eustice, and R. Vasudevan, "Pairwise consistent measurement set maximization for robust multi-robot map merging," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 2916–2923.



JONATHAN M. AITKEN received the M.Eng. and Ph.D. degrees in electronic engineering from the University of York, York, U.K., in 2005 and 2009, respectively. From 2009 to 2013, he worked as a Postdoctoral Researcher with the High Integrity Systems Engineering Group, Department of Computer Science, University of York. He joined the Department of Automatic Control and Systems Engineering, The University of Sheffield, as a Research Fellow, in 2013, working on the develop-

ment of long-term autonomy for robotics in nuclear environments. Since 2018, he has been a University Teacher in Robotics with active research interests in situational awareness, collaborative robotics, and digital twins for robotic systems.



MATHEW H. EVANS received the B.Sc. degree in psychology, the M.Sc. degree in computational and systems neuroscience, and the Ph.D. degree in robotics from The University of Sheffield, Sheffield, U.K., in 2004, 2007, and 2012, respectively. He worked as a Postdoctoral Researcher in bioinspired robotics and computational neuroscience with The University of Sheffield, from 2012 to 2014, and since 2019, The University of Manchester, from 2014 to 2018, and the

University of Nottingham, from 2018 to 2019. He is currently working on developing methods for robot localization and mapping in feature sparse environments.



ROB WORLEY received the M.Eng. degree in integrated mechanical and electrical engineering from the University of Bath, Bath, U.K., in 2018, specializing in mobile robotics, computer vision, machine learning, and control. He is currently pursuing the Ph.D. degree with the Department of Automatic Control and Systems Engineering, The University of Sheffield, studying robot localization for application in buried pipe networks.

SARAH EDWARDS received the M.Sc. degree in robotics from The University of Sheffield, Sheffield, U.K., in 2019, where she is currently pursuing the Ph.D. degree with the Department of Automatic Control and Systems Engineering researching the use of machine learning techniques for automated robotic navigation and exploration.



RUI ZHANG received the M.Sc. degree in autonomous and intelligent systems from The University of Sheffield, Sheffield, U.K., in 2019, specializing in system autonomy and intelligence, machine learning and sensor fusion, where he is currently pursuing the Ph.D. degree with the Department of Automatic Control and Systems Engineering, studying robot localization for applications in buried pipe networks.



TONY DODD received the B.Eng. degree in aerospace systems engineering from the University of Southampton, Southampton, U.K., in 1994, and the Ph.D. degree in machine learning from the Department of Electronics and Computer Science, University of Southampton, in 2000. From 2002 to 2010, he was a Lecturer with the Department of Automatic Control and Systems Engineering, The University of Sheffield, Sheffield, U.K. He was a Senior Lecturer, from 2010 to 2015, and a Professor of Autonomous Systems Engineering, from 2015 to 2019, with The University of Sheffield. Since 2019, he has been an Associate Dean of Research and Enterprise with the School of Digital, Technologies and Arts, Staffordshire University, U.K.



LYUDMILA MIHAYLOVA (Senior Member, IEEE) received the Ph.D. degree. She is currently a Professor of signal processing and control with the Department of Automatic Control and Systems Engineering, The University of Sheffield, Sheffield, U.K. Her research interests include autonomous systems with applications to cities, autonomous and assisted living systems. She received the Tammy Blair Best Award from the International Conference of Information Fusion 2017, the best paper awards from the IEEE DESSERT'2019, 17th IEEE SPA'2013 Conference and IEEE Sensor Data Fusion Workshop, 2013, and others. She is on the Board of Directors of the International Society of Information Fusion (ISIF) and was the ISIF President, during 2016–2018.



SEAN R. ANDERSON received the M.Eng. degree in control systems engineering and the Ph.D. degree in nonlinear system identification from The University of Sheffield, Sheffield, U.K., in 2001 and 2005, respectively.

From 2005 to 2011, he worked as a Postdoctoral Researcher in computational neuroscience, bioinspired robotics and signal processing with The University of Sheffield. He joined the Department of Automatic Control and Systems Engineering, The University of Sheffield, as a Lecturer, in 2012, where he has been a Senior Lecturer, since 2015, working in areas, such as robotics, nonlinear system identification, and computational biology.

...

Chapter 3

Visual Odometry Using Joint and Manholes Detections

This chapter contains a publication detailing the development and testing of an algorithm to perform odometry in sewers using predictable landmarks. The full publication is preceded by a summary of the research presented in the publication.

3.1 Summary of research

The visual odometry system presented here allows a robot to localise itself within a pipe network by leveraging the predictability of the environment (pipe joint spacings) and utilising information that can be obtained above ground prior to operation (manhole locations). While the system cannot achieve the precision of state-of-the-art SLAM algorithms it overcomes many of the challenges that cause those algorithms to perform poorly and often fail entirely in pipe networks (mainly feature sparsity).

Pipe networks generally consist of sections of pipe manufactured to a regular length which are then joined together when placed below ground. These pipes are further connected by manholes to allow human access to the network. While the diameter of pipe sections can vary significantly depending on the pipe locations in the network and intended function, the length of each segment and as such the distance between joints, was found to be relatively consistent in the early stages of research. The total length of pipes, i.e. the distance between manholes, also varies significantly, however, the locations of the manholes can be observed from above

ground. By detecting the joints between pipe sections, the distance travelled along a pipe can be calculated by multiplying the number of joints detected so far by the distance between each joint. Additionally, detecting manholes whose locations have been found prior to operation can be used to reset the position of the robot to a known location and reduce the impact of missed or extraneous joint detections.

The research contained within this publication is primarily concerned with reliably detecting pipe joints in camera footage. Simple methods, such as using the Hough Transform to find circles within each frame, were found to be incapable of identifying joints and joints alone, while initial tests of deep learning algorithms were inconsistent and improvement would come with increased computational complexity and diminishing returns. The primary difficulty faced when detecting landmarks in buried pipes is the lack of distinct features. Pipe walls tend to vary between smooth and featureless or repetitive and indistinct. This causes visual SLAM algorithms to fail due to large sections where insufficient features can be found to operate, and when features are detected they are difficult to match between frames due to their similarity to all other features in the region. Detection methods fail due to the majority of each frame being near identical or essentially noise.

The algorithm detailed below utilises the feature sparsity and regularity by recognising that the vast majority of distinct features in pipe networks are associated with either manholes or pipe joints. By clustering traditional feature descriptors and then analysing the structure and consistency across frames of the ones determined to be pipe joint features, we can detect joints in pipes with an 85% accuracy and an F1 score of 0.78. This shows a stark improvement over the Hough transform method which was tested with a range of parameters and only able to achieve, at most, a 58% accuracy and a 0.11 F1 score. The developed method was found to be sufficient to perform odometry with an average error of 2.50m, significantly lower than that of ORB-SLAM3 (Campos et al., 2021) which failed to operate in all testing environments and achieved an average error of 11.53m when it did work.

Manholes were found to be much simpler to detect within pipes than other landmarks. The publication below uses a support vector machine (SVM) to classify frames as having been taken inside manholes using the same features extracted in the joint detection algorithm. This was 98.5% accurate with an F1 score of 0.98 and allowed the odometry to be further improved to an average error of 1.76m.

The developed algorithms were tested using footage from multiple sewer pipe networks, using experiments conducted in collaboration with the author at iCAIR (The Integrated Civil and Infrastructure Research Centre at the University of Sheffield) and datasets made publicly available by the author at https://orda.shef.ac.uk/articles/dataset/Visual_Odometry_for_Robot_Localisation_in_Feature-Sparse_Sewer_Pipes_Using_Joint_and_Manhole_Detections_---Data/21198070. The novel algorithm is compared to a state-of-the-art visual SLAM algorithm which it was found to outperform considerably.

3.2 Publication



OPEN ACCESS

EDITED BY

Tarveer Saleh,
International Islamic University Malaysia,
Malaysia

REVIEWED BY

Md Raisuddin Khan,
International Islamic University Malaysia,
Malaysia
Tao Huang,
James Cook University, Australia

*CORRESPONDENCE

S. R. Anderson,
✉ s.anderson@sheffield.ac.uk

SPECIALTY SECTION

This article was submitted to Field
Robotics, a section of the journal
Frontiers in Robotics and AI

RECEIVED 24 January 2023

ACCEPTED 16 March 2023

PUBLISHED 06 March 2023

CITATION

Edwards S, Zhang R, Worley R, Mihaylova
L, Aitken J and Anderson SR (2023), A
robust method for approximate visual
robot localization in feature-sparse
sewer pipes.
Front. Robot. AI 10:1150508.
doi: 10.3389/frobt.2023.1150508

COPYRIGHT

© 2023 Edwards, Zhang, Worley,
Mihaylova, Aitken and Anderson. This is
an open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which does
not comply with these terms.

A robust method for approximate visual robot localization in feature-sparse sewer pipes

S. Edwards, R. Zhang, R. Worley, L. Mihaylova, J. Aitken and
S. R. Anderson*

Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield, United
Kingdom

Buried sewer pipe networks present many challenges for robot localization systems, which require non-standard solutions due to the unique nature of these environments: they cannot receive signals from global positioning systems (GPS) and can also lack visual features necessary for standard visual odometry algorithms. In this paper, we exploit the fact that pipe joints are equally spaced and develop a robot localization method based on pipe joint detection that operates in one degree-of-freedom along the pipe length. Pipe joints are detected in visual images from an on-board forward facing (electro-optical) camera using a bag-of-keypoints visual categorization algorithm, which is trained offline by unsupervised learning from images of sewer pipe joints. We augment the pipe joint detection algorithm with drift correction using vision-based manhole recognition. We evaluated the approach using real-world data recorded from three sewer pipes (of lengths 30, 50 and 90 m) and benchmarked against a standard method for visual odometry (ORB-SLAM3), which demonstrated that our proposed method operates more robustly and accurately in these feature-sparse pipes: ORB-SLAM3 completely failed on one tested pipe due to a lack of visual features and gave a mean absolute error in localization of approximately 12%–20% on the other pipes (and regularly lost track of features, having to re-initialize multiple times), whilst our method worked successfully on all tested pipes and gave a mean absolute error in localization of approximately 2%–4%. In summary, our results highlight an important trade-off between modern visual odometry algorithms that have potentially high precision and estimate full six degree-of-freedom pose but are potentially fragile in feature sparse pipes, *versus* simpler, approximate localization methods that operate in one degree-of-freedom along the pipe length that are more robust and can lead to substantial improvements in accuracy.

KEYWORDS

robot localization, sewer pipe networks, feature-sparse, visual odometry, bag-of-keypoints, pipe joint detection

1 Introduction

Sewer networks transport waste products in buried pipes. They are an essential part of our infrastructure but are prone to damage such as cracks, with an estimated 900 billion gallons of untreated sewage discharged into United States of America waterways each year ([American Society of Civil Engineers, 2011](#)). Therefore, sewer pipes need regular monitoring and inspection so that repairs can be effectively targeted and performed. The

traditional way of performing inspection in sewer pipes is *via* manually operated, tethered, CCTV rovers. There is an opportunity to make this process more efficient *via* autonomous robot inspection. One of the key challenges to overcome for this is to solve the robot localization problem so that the location of damage is known.

There are a number of different methods developed for robot localization in pipes (Aitken et al., 2021; Kazeminasab et al., 2021). The methods can be divided based on sensor type: the most simple are dead-reckoning methods based on inertial measurement units (IMUs) and wheel or tether odometry (Murtra and Mirats Tur, 2013; Chen et al., 2019; Al-Masri et al., 2020). The main limitation of these methods is that they drift, and so some authors have introduced drift correction methods based on known landmarks including pipe joints (where accelerometers are used to detect the vibration as the robot moves over the joint) (Sahli and El-Sheimy, 2016; Guan et al., 2018; Wu et al., 2019; Al-Masri et al., 2020), and above-ground reference stations (Wu et al., 2015; Chowdhury and Abdel-Hafez, 2016).

Cameras are another widely-used method of localization in pipe robots, using monocular visual odometry (VO) (Hansen et al., 2011a; Hansen et al., 2013; Hansen et al., 2015), visual simultaneous localization and mapping (vSLAM) (Evans et al., 2021; Zhang et al., 2021), stereo VO (Hansen et al., 2011b), and RGB-D cameras (Alejo et al., 2017; Alejo et al., 2019). Laser scanners have been used in pipes for recognising landmarks such as manholes, junctions and elbows (Ahrary et al., 2006; Lee et al., 2016; Kim et al., 2018) although not, it would appear, for the odometry problem. Finally, acoustic and radio frequency (RF) signals such as ultrasonic (Ma et al., 2015), hydrophone (Ma et al., 2017a; Ma et al., 2017b; Worley et al., 2020a), RF (Seco et al., 2016; Rizzo et al., 2021), low frequency acoustic (Bando et al., 2016) and acoustic-echo (Worley et al., 2020b; Yu et al., 2023) methods have been used, but these are still emerging technologies.

The sensor technology that is most of interest in this paper for localization is cameras. This is because sewer pipe inspection is often conducted using vision-based methods (Duran et al., 2002; Myrans et al., 2018) and most pipe inspection robots developed to date include cameras for visual inspection, e.g., MAKRO (Rome et al., 1999), KANTARO (Nassiraei et al., 2006), MRINSPECT (Roh et al., 2008), PipeTron (Debenest et al., 2014), EXPLORER (Schempf et al., 2010) and recent miniaturized pipe inspection robots (Nguyen et al., 2022). Therefore, it is appealing to make dual use of a camera for both inspection and localization.

The main challenge facing camera-based localization in pipes is that standard visual odometry algorithms for localization based on keyframe optimisation methods, e.g., Hansen et al. (2011a); Hansen et al. (2013, 2015); Zhang et al. (2021); Evans et al. (2021) tend to fail in environments that lack visual features, and this is particularly the case for newer sewer pipes, although we have shown in aged sewer pipes that sufficient features exist for these methods to work well (Evans et al., 2021). We will go on to show in the results that a standard feature-based keyframe optimisation method for visual SLAM, ORB-SLAM3 (Campos et al., 2021), fails in these types of feature-sparse sewer pipe. There is a key research gap, therefore, in developing a visual odometry method for sewer pipes that lack visual features, which is the problem that we address here.

In this paper, we propose a new solution to the problem of robot localization in feature-sparse sewer pipes based on joint and manhole detections using camera images. Joint detection can be used for localization because joints occur at regularly spaced intervals where the inter-joint distance can be known *a priori* from installation data records, or estimated from odometry. The robot location along a pipe length, in one-degree of freedom, can be approximately obtained from scaling the count of pipe joints by the inter-joint distance. This transforms the problem of robot localization to one of pipe joint detection. Manholes can be mapped from above-ground to serve as drift-correcting landmarks when detected from inside the pipe to further improve the localization system.

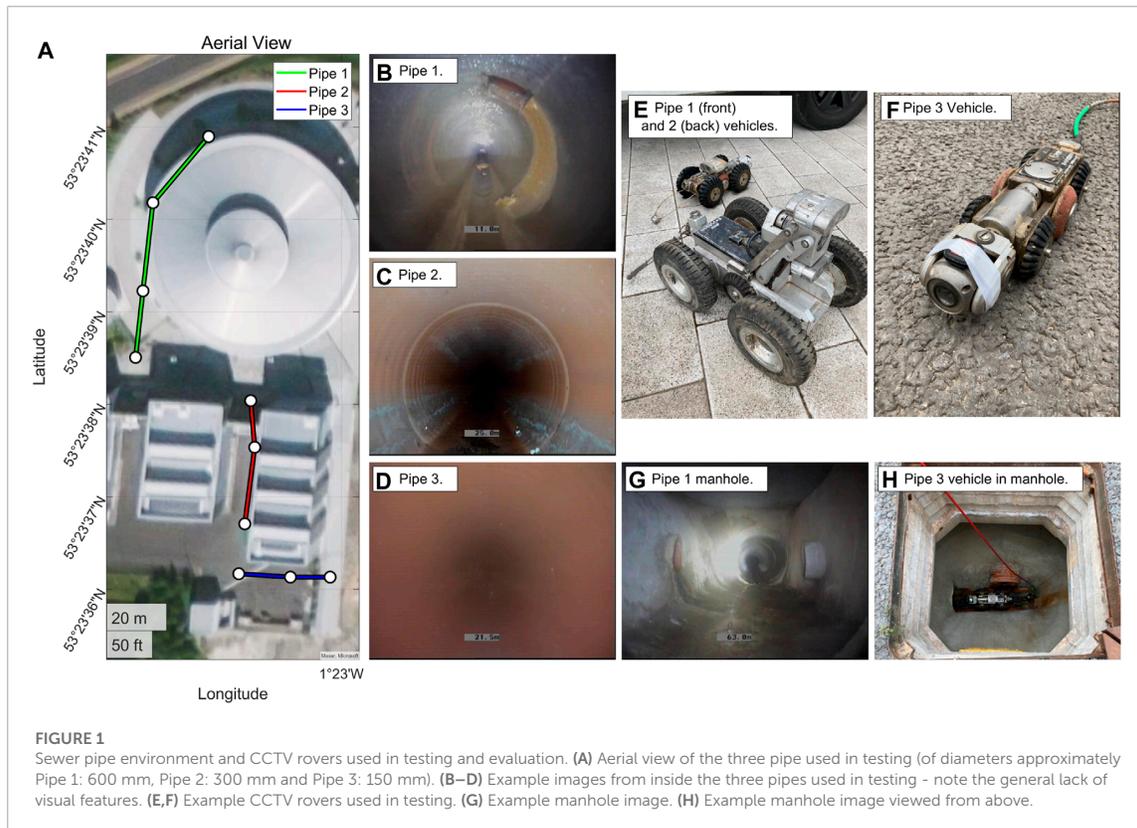
Joint detection in pipes has been previously addressed using vision methods for the purpose of damage detection (not localization), using forward facing cameras (Pan et al., 1995), omnidirectional cameras (Matsui et al., 2010; Mateos and Vincze, 2011) and fusion of laser scanners with cameras (Kolesnik and Barotoff, 2000). For forward facing cameras, the standard approach to pipe joint detection is to apply circle detection using the Hough transform to each image frame (Pan et al., 1995). However, in exploratory analysis we found the Hough transform approach was unreliable, often detecting spurious circles. Instead, here we use SURF for image feature extraction (Bay et al., 2008), followed by feature selection for pipe joints using a bag-of-keypoints method (Csurka et al., 2004), followed by circle fitting to detect the joint. The key advantage of our approach is that it enables us to train the feature selection algorithm on representative examples of sewer pipe joints (in an unsupervised manner), whilst the Hough transform does not have access to this prior information that specializes the method to the sewer pipe environment. We make our procedure even more robust by performing joint detections across a window of frames.

A potential limitation of only using pipe joints for localization is that detection errors can be made, such as a joint being missed (a false negative), or counted when not present (a false positive). To address this challenge, we develop a modular extension based on manhole detection to correct for drift in the localization algorithm: we assume the manhole locations are known or can be mapped from above-ground offline, and then use a linear classifier to detect the manholes from within the pipe. We use the *same* SURF image features as input to both the joint and manhole detection systems, making the approach more computationally efficient than using completely separate systems. Overall, the joint detection with manhole drift correction produces an accurate and robust method for localization.

To test and evaluate the method, we use real sewer pipe data taken from three different types of pipe to demonstrate its effectiveness (Figure 1), data available at The University of Sheffield data repository ORDA https://figshare.shef.ac.uk/articles/dataset/Visual_Odometry_for_Robot_Localisation_in_Feature-Sparse_Sewer_Pipes_Using_Joint_and_Manhole_Detections_-_Data/21198070. We benchmark against ORB-SLAM3 (Campos et al., 2021) as a standard method for visual odometry.

In summary, the main contributions of the paper are as follows.

- A robust vision-based method for approximately localizing a robot (to the nearest pipe joint) along the lengths of feature-sparse sewer pipes using joint detections combined with a



method for drift correction at manhole locations using vision-based automated manhole detection.

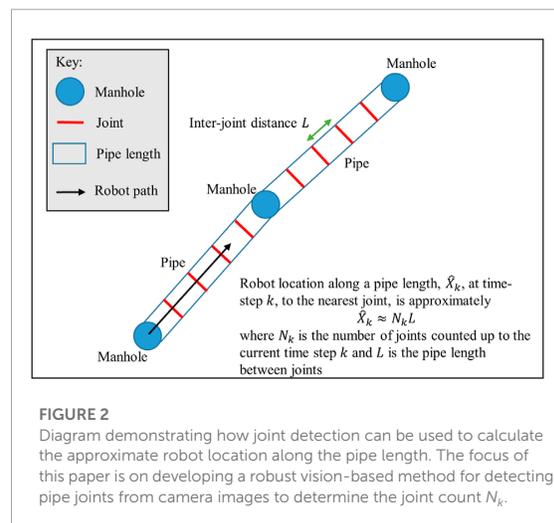
- A method for robustly detecting joints in pipes using a bag-of-keypoints visual categorization algorithm that is benchmarked against a standard method for detecting pipe joints in images - the Hough transform.
- Experimental testing and evaluation of the localization method using real-data gathered from three live sewer pipes.
- Benchmarking of the localization method against a well-known, state-of-the-art visual SLAM algorithm - ORB-SLAM3 (Campos et al., 2021).

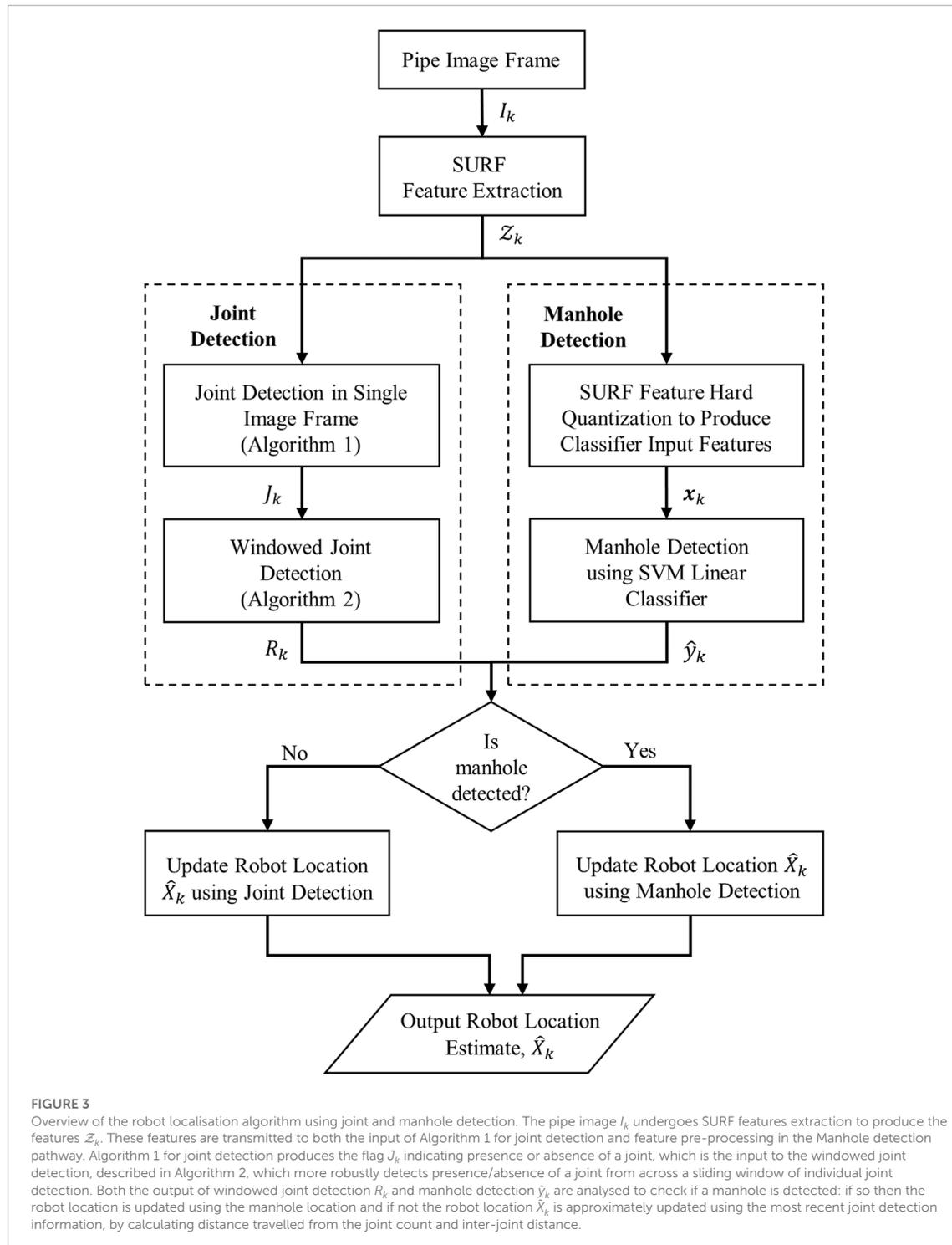
The paper is structured as follows. In Section 2 we describe the methods and particularly our new algorithm for localization using pipe joint detections and manhole detections, as well as the dataset for evaluation. In Section 3 we give the results of the algorithm on real-world sewer pipe data and include a benchmark comparison to ORB-SLAM3. In Section 4 we provide a discussion and in Section 5 we summarise the main achievements of the paper.

2 Methods

In this section we describe the joint detection algorithm, manhole detection and the experimental data collection used to

evaluate the algorithm in a real-world live sewer pipe. The link between vision-based joint detection and robot localization along a pipe length is described in Figures 2, 3 gives an overview of the





methods used for vision-based joint and manhole detection, and robot localization.

2.1 Joint detection

The robot location along the pipe, \hat{X} , i.e., in one degree of freedom, can be obtained from joint detections because joints occur at regularly spaced intervals, where we assume the inter-joint distance is known *a priori* or we assume can be estimated. The joint detection algorithm proposed here has a few main steps described below: feature extraction, circle fitting and windowed joint detection.

The purpose of the feature extraction step in the algorithm is to find points of interest within the current test image frame I_k at time step k , to check for a joint. In this paper we use a method inspired by visual categorization with bags of keypoints (Csurka et al., 2004), because it is simple, fast and effective and therefore well suited to small, low-powered robots for the pipe environment. The method operates by first extracting features from a test image I_k using speeded up robust features (SURF) (Bay et al., 2008),

$$\mathcal{Z}_k = \{\mathbf{z}_{k,1}, \dots, \mathbf{z}_{k,n}\} \quad (1)$$

Where $\mathbf{z}_{k,i} \in \mathbb{R}^d$ is a SURF feature of dimension d , and n is the total number of SURF features found in an image. We then use a threshold test to retain only those features sufficiently similar to keypoints previously seen in joints in training data,

$$\mathcal{Z}_k^* = \{\mathbf{z}_{k,i} : d(\mathbf{z}_{k,i}, \mathbf{k}_j) \leq \beta\} \quad \forall i, j \quad (2)$$

Where $d(\mathbf{z}_{k,i}, \mathbf{k}_j)$ is a distance-metric (Euclidean in this case) of feature $\mathbf{z}_{k,i}$ from keypoint \mathbf{k}_j and β is a threshold parameter tuned offline. The keypoints, \mathbf{k}_j , are obtained offline by using a set of training data with a K-means clustering algorithm, similar to (Csurka et al., 2004). The use of multiple clusters enables the method to be robust to the variation in appearance of pipe joints, whilst also excluding non-joint features from detection. Identifying a joint does not require a large number of features, therefore it is more important to exclude false detections than to detect every relevant feature, so the threshold, β for selection is tuned to be more exclusive than inclusive. Additionally, points lying within regions of an image that

are known to not contain joints, such as the centre, can be excluded automatically using a mask, M .

We perform the actual joint detection by fitting a circle to the extracted features, \mathcal{Z}_k^* , to test whether the features resemble a joint. This is justified because it is known that sewer pipes are cylindrical, and the position of the robot's camera is relatively central in the pipe during normal forward motion. As such, the pipe joints will appear close to circular in images captured by robots in pipes. Note that we only check for a joint if the number of features in \mathcal{Z}_k^* is greater than a threshold parameter, γ . Note also that the objective here is to detect the pipe joint, not obtain an optimal model of a circle, therefore we avoid computationally intensive circle fitting based on iterative optimisation (Gander et al., 1994), and opt for a more rapid and simple estimate: we take the average of all feature points as the circle centre, \mathbf{c}_k , and the average distance from the centre of each point as the radius, r_k ,

$$\mathbf{c}_k = (\bar{u}, \bar{v}) \quad (3)$$

$$r_k = \frac{1}{N} \sum_i^N \sqrt{(u_i - \bar{u})^2 + (v_i - \bar{v})^2} \quad (4)$$

where (u_i, v_i) is the horizontal-vertical image coordinates corresponding to feature $\mathbf{z}_{k,i}$, \bar{u} is the mean of the u -coordinates and \bar{v} is the mean of the v -coordinates. This circle fitting method is computationally efficient but its disadvantage is that it will fit the circle to any set of points it is given, regardless of the actual geometry of the points (Gander et al., 1994). To compensate for this, the joint is only detected if the centre and radius of the fitted circle are within predefined threshold parameters δ_1 and δ_2 respectively. This procedure for detecting a joint in a single image frame I_k is described in Algorithm 1 (Figure 4).

The procedure for detecting a joint is made more robust here by performing detections across a window of frames. This reduces the impact of false detections, whilst also providing robustness against multiple correct, but discontinuous, detections of a single joint. To raise a joint detection flag, $R_k = 1$, the method simply requires that the number of positive joint detections from Algorithm 1, over a window of frames of size n_w , is greater than a threshold parameter ζ and that the variances of the centre and radius of the detected joints are below thresholds, $\sigma_c^2 \leq \eta_1$ and $\sigma_r^2 \leq \eta_2$. The reverse procedure is used to lower the joint detection flag, $R_k = 0$. This procedure is described in Algorithm 2 (Figure 5).

Algorithm 1: jointDetection

Input: SURF features $\mathcal{Z}_k = \{\mathbf{z}_{k,1}, \dots, \mathbf{z}_{k,n}\}$ from image I_k

Output: joint detection boolean variable J_k at time step k

- 1: eliminate SURF features $\mathbf{z}_{k,i}$ in the pipe centre using mask M
- 2: select features within threshold distance to K-means cluster centres, $\mathcal{Z}_k^* = \{\mathbf{z}_{k,i} : d(\mathbf{z}_{k,i}, \mathbf{k}_j) \leq \beta\}$
- 3: **if** sufficient features are selected, $|\mathcal{Z}_k^*| \geq \gamma$ **then**
- 4: fit circle to image coordinates (u_i, v_i) corresponding to features $\mathbf{z}_{k,i} \in \mathcal{Z}_k^*$, with circle centre \mathbf{c}_k and radius r_k
- 5: **end if**
- 6: **if** the centre or radius of the circle are outside threshold values, $c_k \geq \delta_1$ or $r_k \geq \delta_2$, **then**
- 7: record no joint detected for image I_k , set $J_k = 0$
- 8: **else**
- 9: record joint detected for image I_k , set $J_k = 1$
- 10: **end if**

FIGURE 4

Algorithm 1: Joint detection in a single image frame.

Algorithm 2: windowedJointDetection

Input: joint detections from Algorithm 1 over a window n_w of image frames, J_{k-n_w+1}, \dots, J_k and current state of robot R_{k-1}

Output: windowed joint detection boolean variable R_k at time step k

```

1: if the robot is not currently considered to be at a joint,  $R_{k-1} = 0$ , then
2:   if joint detections in a window of frames satisfy  $\sum_{i=0}^{n_w-1} J_{k-i} > \zeta$ , and,  $\sigma_c^2 \leq \eta_1$  and  $\sigma_r^2 \leq \eta_2$ , then
3:     set the robot state to be at a joint,  $R_k = 1$ 
4:   else
5:     set the robot state to not be at a joint,  $R_k = 0$ 
6:   end if
7: else if the robot is currently considered to be at a joint,  $R_{k-1} = 1$ , then
8:   if joint detections in a window of frames satisfy  $\sum_{i=0}^{n_w-1} J_{k-i} \leq \kappa$  or  $\sigma_c^2 \geq \lambda_1$  and  $\sigma_r^2 \geq \lambda_2$  then
9:     set the robot state to not be at a joint,  $R_k = 0$ 
10:  else
11:    set the robot state to be at a joint,  $R_k = 1$ 
12:  end if
13: end if
    
```

FIGURE 5

Algorithm 2: Detecting a joint in a window of frames.

All parameters in Algorithms 1 and 2 were tuned using a grid search across real sewer pipe data (except $\gamma, \delta_1, \delta_2$ and ζ that were tuned manually). The grid search took a number of days to evaluate because it involved a relatively long distance with a large number of parameters ($\beta, \eta_1, \eta_2, n_w, \lambda_1, \lambda_2$) and although reaching a global optimum could not be guaranteed, as a global search method a grid search is relatively robust to not well-distinguishable local optima, and testing on independent validation data ensured good generalization.

The robot location \hat{X}_k along a pipe, can be obtained relative to a starting position by using the joint detections, R_k , to count the number of joints detected up to the current time step, N_k , and then taking the product of the joint count with the inter-joint distance, L ,

$$\hat{X}_k \approx N_k L. \quad (5)$$

2.2 Manhole detection

Manhole detections can be used to correct drift in the joint localization algorithm. Manhole locations can be known *a priori* or mapped from above-ground, and then detected from within the pipe. While many potential methods of detecting manholes exist, we continue to rely on camera data and use a bag-of-features image recognition system to detect manholes. The advantage of this approach is that the *same* SURF features extracted for joint detections, Z_k , can be used for manhole detection. This makes the approach more computationally efficient than using a separate feature extraction/classification method.

To detect manholes, we define a standard binary classification problem of manhole *versus* no manhole, using a linear support vector machine (SVM). We construct a visual vocabulary of features offline along with clusters using K-means clustering (Wang and Huang, 2015), and then in online operation the SURF features, z_k , extracted from image frame I_k , undergo hard quantization by representing each local feature by the nearest

visual word, which produces the classifier input features \mathbf{x}_k . The classifier training/validation dataset of input-output pairs is therefore represented as

$$\mathcal{D} = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_m, y_m)\}, \quad (6)$$

where the binary output $y_k \in \{+1, -1\}$, represents the classes manhole and no manhole respectively. We use a standard linear (soft margin) SVM to classify the presence of a manhole, where the decision hyperplane is defined as

$$f(\mathbf{x}_k) = \mathbf{w}^T \mathbf{x}_k + b = 0, \quad (7)$$

where \mathbf{w} and b are the parameters that define the decision hyperplane and the classifier predicts the class label using the sign of $f(\mathbf{x}_k)$. Optimal parameters for the soft-margin SVM were estimated here using sequential minimal optimization (Cervantes et al., 2020) using a balanced dataset of sample images taken from pipes 1, 2 and 3, and performance was evaluated using 10-fold cross-validation.

In online operation detections of manholes are windowed similarly to joints (as in Figure 3) in order to limit the impact of single errant frames. Finally, the robot location, \hat{X}_k , can be determined from the any detected manhole because manhole locations are assumed known.

2.3 Experimental data

Tests were conducted on data gathered from real-world, live, buried sewer pipes at The Integrated Civil and Infrastructure Research Centre (iCAIR), at The University of Sheffield, UK, using tethered mobile CCTV platforms commonly used in pipe inspections. As can be seen in Figure 1, three different platforms were used to allow a variety of pipes to be tested.

The pipe lengths were approximately 90 m for pipe 1, 50 m for pipe 2 and 30 m for pipe 3. Diameters were approximately 600 mm for pipe 1, 300 mm for pipe 2 and 150 mm for pipe 3.

The data used to train both the joint detection and manhole detection systems was selected in part from the data used to test the systems, but also from data in other pipes at the same location and collected in the same experiment.

Ground truth localization was obtained from the robot tether, which measured distance travelled.

2.4 Testing and evaluation

The localization system proposed here was compared and benchmarked against a standard method for visual odometry, ORB-SLAM3. To do this, the camera intrinsics were obtained from a standard checkerboard calibration procedure.

To evaluate the overall effectiveness of the robot localization system using our proposed method and ORB-SLAM3 we used the mean absolute error in localization,

$$MAE = \frac{1}{n} \sum_{i=1}^n |X_i - \hat{X}_i|, \tag{8}$$

where X_i is the true robot location in terms of distance travelled along the pipe map (as measured using the tether on the robot) and \hat{X}_i is the estimated robot location.

We also used the percent of operating time spent with an error below a threshold,

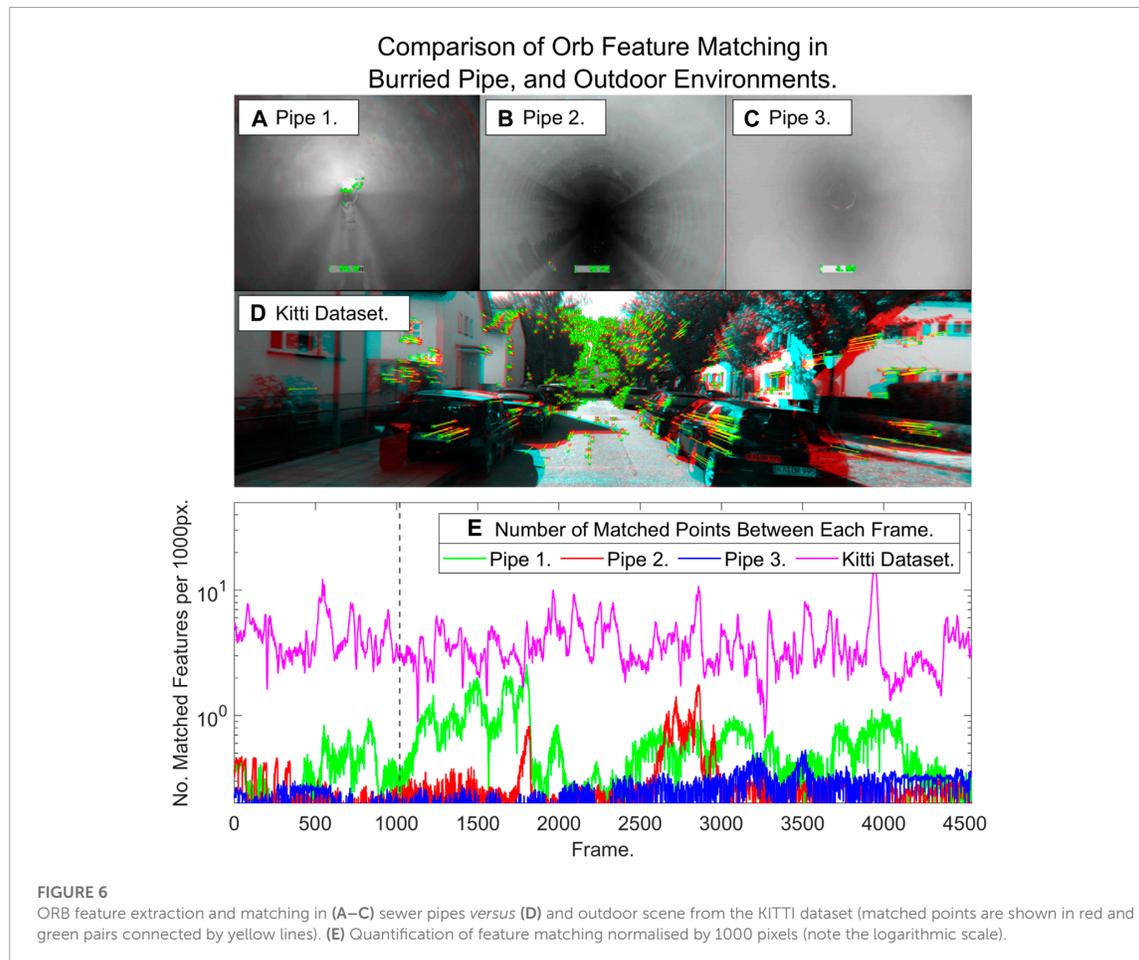
$$E_{\%} = \frac{100}{n} \sum_{i=1}^n g(X_i), \tag{9}$$

where

$$g(X_i, \hat{X}_i) = \begin{cases} 1 & |X_i - \hat{X}_i| \leq E_{\text{thresh}} \\ 0 & |X_i - \hat{X}_i| > E_{\text{thresh}} \end{cases} \tag{10}$$

and where E_{thresh} is the known distance between each pipe joint, which would be the maximum error if the system was performing ideally.

To evaluate the manhole classifier and joint detections we used the metrics accuracy A (manholes only), recall R , precision P and F1



score F_1 (Powers, 2011).

$$A = \frac{TP + TN}{TP + TN + FP + FN}, \tag{11}$$

$$P = \frac{TP}{TP + FP}, \tag{12}$$

$$R = \frac{TP}{TP + FN}, \tag{13}$$

$$F_1 = 2 \frac{P \cdot R}{P + R}, \tag{14}$$

Where TP is true positive, TN is true negatives, FP is false positives and FN is false negatives.

3 Results

3.1 Feature detection in sewer pipes vs urban environments

In this section we analyse the nature of the sewer pipe environment with regard to the prevalence of features and compare to outdoor urban environments where visual odometry is often applied. We compare ORB feature extraction and matching in our sewer pipes to a sequence from the well-known KITTI dataset (Geiger et al., 2013): as can be seen in Figure 6A–C versus

TABLE 1 Comparison of joint detection using the bag-of-keypoints method versus the Hough transform (where the edge threshold and sensitivity of the Hough transform are varied systematically). Note that in certain instances the Hough transform fails to detect any circles therefore Precision and F1-score are undefined, which is indicated with a “-”.

	Joint detection with bag-of-keypoints	Hough transform							
Edge Thresh	-	0	0.25	0	0.25	0	0.25	0	0.25
Sensitivity	-	0.85	0.85	0.9	0.9	0.95	0.95	1	1
Accuracy	84.65%	57.35%	58.01%	24.67%	58.01%	5.38%	58.01%	5.77%	0.66%
Precision	0.85	0.00	-	0.08	-	0.05	-	0.06	0.01
Recall	0.73	0.00	0.00	0.16	0.00	0.14	0.00	0.16	0.02
F1-score	0.78	-	-	0.11	-	0.08	-	0.08	0.01

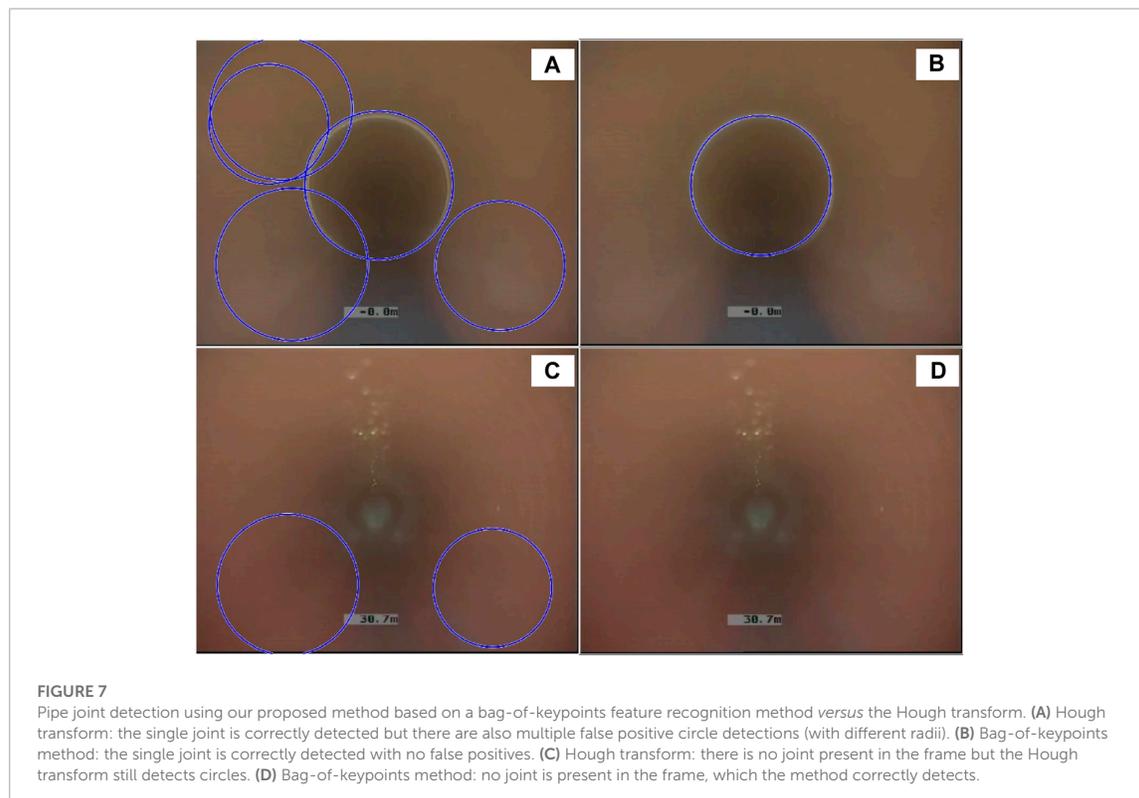


Figure 6D, images from pipe interiors often contain significantly less regions of high texture than those in outdoor environments where visual SLAM systems are known to work well. **Figure 6E** shows that the number of features matched often an order of magnitude lower than in an outdoor environment. This causes frame by frame feature matching algorithms to frequently fail in sewer pipes.

3.2 Joint detection

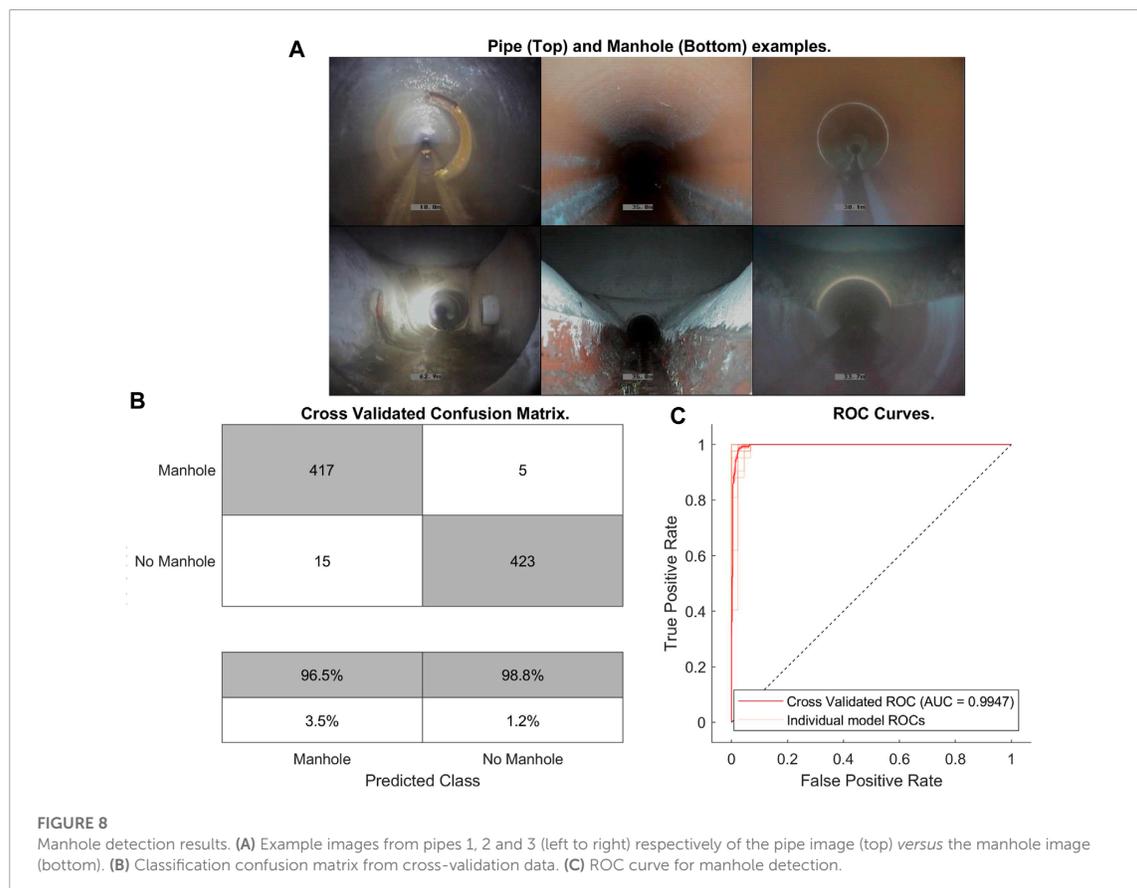
In this section, we provide the results of our proposed joint detection algorithm (Algorithm 1 only, i.e., no windowing) along with a comparison to the Hough transform, which is a standard method of circle detection in images (Yuen et al., 1990), as used in the Matlab function *imfindcircles* and the OpenCV function *cv2.HoughCircles*. To evaluate and compare the methods we used all image frames from pipe 2 and varied the Hough transform tuning parameters, edge threshold and sensitivity, systematically to search out the best performance. The results demonstrate the problems with using the Hough transform compared to our method (**Table 1**) - the edge threshold parameter requires tuning to a value of zero to detect any circles, which leads to a large number of false positives. Consequently, the accuracy is generally low, approximately 25% using the Hough transform (with edge threshold zero and sensitivity

0.9), mainly due to the detection of large numbers of false positives, compared to 85%, using our proposed method of joint detection (**Figure 7A–D**).

3.3 Manhole detection

In this section we provide results of manhole detections using linear SVM classification. Manhole detection performed well, with an accuracy on 10-fold cross validation of 98.5% (with precision 0.97, recall 0.99 and F1-score 0.98): **Figure 8A** shows examples of pipe and manhole environments which illustrates the high visual difference between the two, as well as the spatial distinctions which may be exploitable by other sensors. 10-fold cross validation was performed on a selection of images from pipes and manholes across a variety of pipes.

The network displays a high accuracy as can be seen in **Figure 8B** as well as in **Figure 8C** which also indicates both a high specificity and sensitivity. The shown accuracy is sufficient for the system to reliably detect every manhole it encounters, and a windowing method similar to that used in the joint detection algorithm prevents any false positives or negatives from causing the same manhole to be detected twice.



3.4 Localization

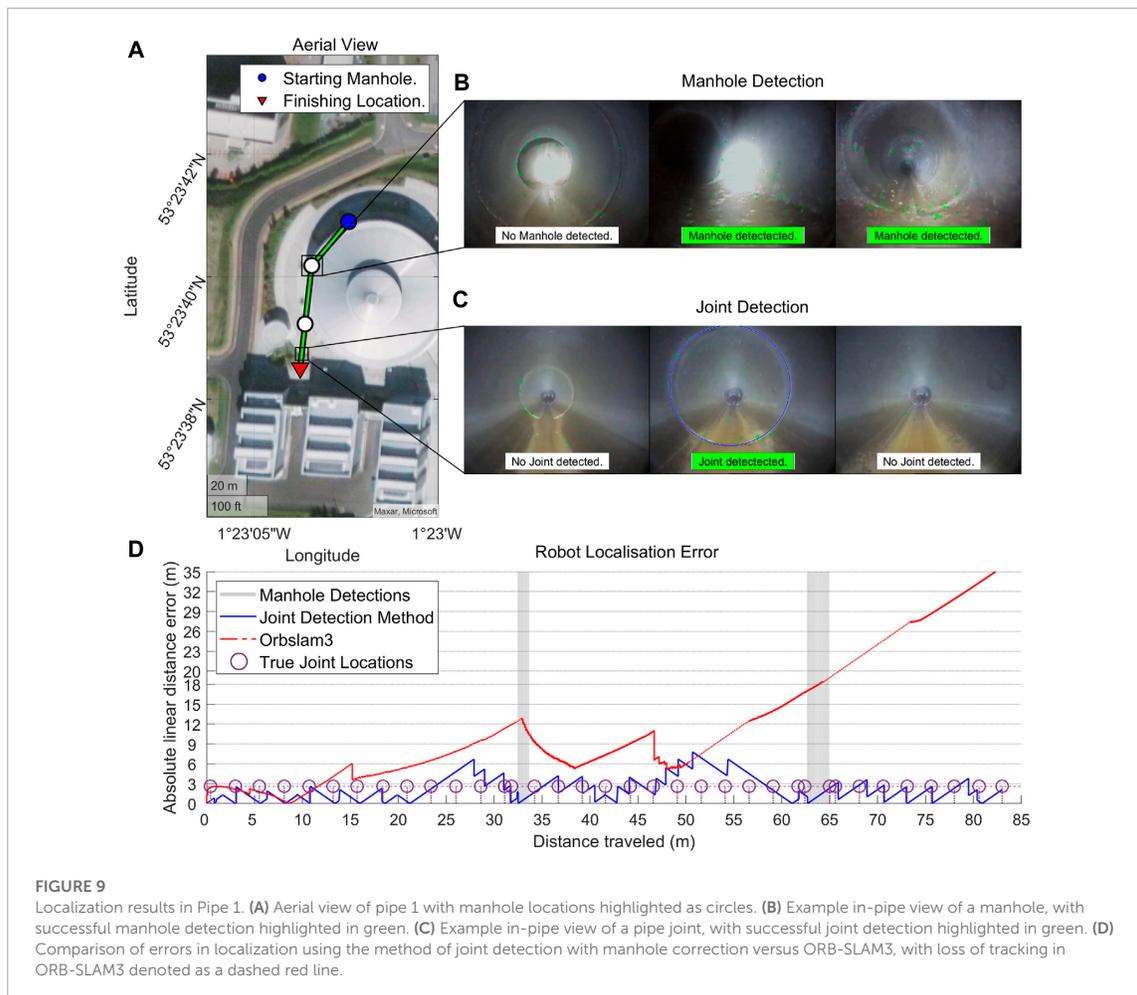
Localization accuracy using the proposed joint detection algorithm improved substantially over using ORB-SLAM3: the average mean absolute error for the joint detection algorithm was 1.8 metres, whilst for ORB-SLAM3 it was 11.5 m in pipes 1 and 2 and a complete failure in pipe 3 due to lack of features. Joint detection, however, worked well in all pipes tested with an F1-score of 0.72–0.95 (Table 2). It is worth also emphasising that although ORB-SLAM3 produced a result in

pipes 1 and 2, it still frequently lost feature tracking and re-initialised due to lack of feature matching. Figure 9 provides a detailed illustration and comparison of localization methods in pipe 1.

Table 2 also shows measured error metrics from the Joint Odometry, both with and without manhole correction, and compares them to ORB-SLAM3. The localization results using ORB-SLAM3 correspond to a mean absolute error of approximately 12%–20% on pipes 1 and 2 (failure in pipe 3), whilst our method worked successfully on all tested pipes and gave a mean absolute

TABLE 2 Localization results from pipe joint detection versus ORB-SLAM3 from three sewer pipes.

Pipe	Joint detection results			Localization accuracy					
	Recall	Precision	F1 Score	Mean Absolute Error (m)			Time Spent Under Target Error (%)		
				Joint + Manhole	Joint only	ORB-SLAM3	Joint + Manhole	Joint only	ORB-SLAM3
Pipe 1	0.69	0.76	0.72	2.04	2.25	12.70	70.78	69.15	11.52
Pipe 2	1	0.80	0.89	2.11	4.23	10.36	72.90	37.30	15.38
Pipe 3	1	0.91	0.95	1.12	1.03	N/A	97.53	97.53	N/A



error in localization of approximately 2%–4% across all pipes, which was a substantial improvement.

4 Discussion

4.1 Overall performance

The aim of this paper was to develop a localization system for feature-sparse sewer pipes based on visual joint and manhole detection, to overcome the limitations of conventional keyframe optimisation visual odometry systems. The results have demonstrated that this objective was successfully achieved. While the system lacks the precision often desired for odometry systems due to its discrete nature, it is able to perform highly accurate localization given relatively limited knowledge of the operating environment. Additionally, discrete updates based on prior external information free the system from problems such as scale ambiguity and loss of tracking that are particularly difficult to overcome in pipe environments. Finally, while still present in the system, errors are accumulated every distance update rather than every frame and are smaller relative to the update than in traditional visual odometry systems, meaning that drift cannot accumulate quick enough to cause system failure before manhole detection corrects the state estimate.

The systems main cause of failure is lower precision and recall in more varied pipe environments, however it should be noted that the systems parameters can be optimised to improve performance in a single pipe at the expense of others.

4.2 Future work

In future work, a number of improvements could be investigated to the system presented here. The first improvement would be an online adaptive method for adjusting the joint detection algorithm parameters automatically while in operation, to account for minor differences between pipes.

The second improvement would be to automatically estimate inter-joint distances. Here we assume these distances are known *a priori*, which is realistic for some pipes. However, this knowledge might not always be available. We have found that simple odometry methods, such as wheel odometry, may be accurate enough over the short distances between pipe joints to derive this information during operation. Alternatively, it should be possible to use the detection of manholes, mapped from above-ground, to estimate the inter-joint distances. These methods require development and testing in future work. In addition, future work could address specific problems where joints are irregularly spaced and pipe bends occur - the latter problem could be addressed by combining joint detection for measuring distance travelled with an IMU to sense changes of direction.

Thirdly, false positives are primarily associated with manholes, however, other predictable environmental features, such as connecting pipes, are also known to reduce the accuracy of the detection system. These other predictable features provide

opportunity for further improvement through their detection or further exploitation of prior knowledge.

5 Summary

In this paper we developed a localization method for sewer pipe inspection robots, operating in pipes with sparse visual features. The method exploited the intrinsic characteristic of the sewer pipe environment, that pipe joints occur at regularly spaced intervals. Therefore the localization problem was transformed to one of pipe joint detection. To further robustify the procedure, manhole detection was also included, which enabled drift correction based on manholes that could be mapped from above-ground. The visual localization algorithm was evaluated on three different real-world, live sewer pipes and then benchmarked against a standard method for visual odometry - ORB-SLAM3. We showed that our method substantially improved on the accuracy and robustness of ORB-SLAM3. Whilst visual SLAM algorithms such as ORB-SLAM3 are sophisticated, potentially very accurate and estimate the full six degree-of-freedom robot pose compared to our discrete, approximate localization method that only works in one-dimension along the pipe length, we would note that there is a trade-off in accuracy and robustness here, with ORB-SLAM3 regularly failing to track features in these feature-sparse sewer pipes. Ultimately we might find both systems are used in parallel in future to take advantage of the attributes of both approaches. The developed method can be applied as part of real robot localization systems, as part of digital twins for pipe networks with different scale, under different environmental conditions and different prior knowledge.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://figshare.shef.ac.uk/articles/dataset/Visual_Odometry_for_Robot_Localisation_in_Feature-Sparse_Sewer_Pipes_Using_Joint_and_Manhole_Detections_-_Data/21198070.

Author contributions

SE, JA, and SRA jointly conceived the approach to robot localization; SE implemented the methods and obtained the results; all authors analyzed the results; SE and SRA wrote the manuscript; all authors commented on and edited the manuscript.

Funding

The authors gratefully acknowledge the support of EPSRC UK funding Council Grant, Pervasive Sensing

for Buried Pipes (Pipebots), Grant/Award Number: EP/S016813/1.

Acknowledgments

The authors acknowledge the assistance of M. Evans, W. Shepherd and S. Tait in organising the data collection.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships

References

- Ahrary, A., Kawamura, Y., and Ishikawa, M. (2006). "A laser scanner for landmark detection with the sewer inspection robot KANTARO," in Proceedings of the IEEE/SMC International Conference on System of Systems Engineering, Los Angeles, CA, USA, 24-26 April 2006 (IEEE), 291–296.
- Aitken, J. M., Evans, M. H., Worley, R., Edwards, S., Zhang, R., Dodd, T., et al. (2021). Simultaneous localization and mapping for inspection robots in water and sewer pipe networks: A review. *IEEE Access* 9, 140173–140198. doi:10.1109/access.2021.3115981
- Al-Masri, W. M. F., Abdel-Hafez, M. F., and Jaradat, M. A. (2020). Inertial navigation system of pipeline inspection gauge. *IEEE Trans. Control Syst. Technol.* 28, 609–616. doi:10.1109/tcst.2018.2879628
- Alejo, D., Caballero, F., and Merino, L. (2019). A robust localization system for inspection robots in sewer networks. *Sensors* 19, 4946. doi:10.3390/s19224946
- Alejo, D., Caballero, F., and Merino, L. (2017). "RGBD-based robot localization in sewer networks," in Proceedings of the IEEE International Conference on Intelligent Robots and Systems, Vancouver, BC, Canada, 24-28 September 2017 (IEEE), 4070–4076.
- American Society of Civil Engineers (2011). Failure to act: The economic impact of current investment trends in water and wastewater treatment infrastructure. Available at: http://www.asce.org/uploadedfiles/issues_and_advocacy/our_initiatives/infrastructure/content_pieces/failure-to-act-waterwastewater-report.pdf (accessed 09 29, 2020).
- Bando, Y., Suhara, H., Tanaka, M., Kamegawa, T., Itoyama, K., Yoshii, K., et al. (2016). "Sound-based online localization for an in-pipe snake robot," in Proceedings of the International Symposium on Safety, Security and Rescue Robotics, Switzerland, 23-27 October 2016 (Institute of Electrical and Electronics Engineers Inc.), 207–213.
- Bay, H., Ess, A., Tuytelaars, T., and Van Gool, L. (2008). Speeded-up robust features (surf). *Comput. Vis. Image Underst.* 110, 346–359. doi:10.1016/j.cviu.2007.09.014
- Campos, C., Elvira, R., Rodríguez, J. J. G., Montiel, M., and Tardós, D. (2021). ORB-SLAM3: An accurate open-source library for visual, visual-inertial, and multimodal SLAM. *IEEE Trans. Robotics* 37, 1874–1890. doi:10.1109/tro.2021.3075644
- Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L., and Lopez, A. (2020). A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing* 408, 189–215. doi:10.1016/j.neucom.2019.10.118
- Chen, Q., Zhang, Q., Niu, X., and Wang, Y. (2019). Positioning accuracy of a pipeline surveying system based on MEMS IMU and odometer: Case study. *IEEE Access* 7, 104453–104461. doi:10.1109/access.2019.2931748
- Chowdhury, M. S., and Abdel-Hafez, M. F. (2016). Pipeline inspection gauge position estimation using inertial measurement unit, odometer, and a set of reference stations. *ASME J. Risk Uncertain. Part B* 2, 21001. doi:10.1115/1.4030945
- Csurka, G., Dance, C., Fan, L., Willamowski, J., and Bray, C. (2004). "Visual categorization with bags of keypoints," in *Workshop on statistical learning in computer vision* (Prague, Czech Republic): ECCV).
- Debenest, P., Guarnieri, M., and Hirose, S. (2014). "PipeTron series - robots for pipe inspection," in Proceedings of the 3rd International Conference on Applied Robotics for the Power Industry (CARPI), Brazil, 14-16 October 2014 (IEEE), 1–6.
- Duran, O., Althoefer, K., and Seneviratne, L. D. (2002). State of the art in sensor technologies for sewer inspection. *IEEE Sensors J.* 2, 73–81. doi:10.1109/jsen.2002.1000245
- Evans, M. H., Aitken, J. M., and Anderson, S. R. (2021). "Assessing the feasibility of monocular visual simultaneous localization and mapping for live sewer pipes: A field robotics study," in Proc. of the 20th International Conference on Advanced Robotics (ICAR), Slovenia, 06-10 December 2021 (IEEE), 1073–1078.
- Gander, W., Golub, G. H., and Strelbel, R. (1994). Least-squares fitting of circles and ellipses. *BIT Numer. Math.* 34, 558–578. doi:10.1007/bf01934268
- Geiger, A., Lenz, P., Stiller, C., and Urtasun, R. (2013). Vision meets robotics: The KITTI dataset. *Int. J. Robotics Res.* 32, 1231–1237. doi:10.1177/0278364913491297
- Guan, L., Gao, Y., Noureldin, A., and Cong, X. (2018). "Junction detection based on CCWT and MEMS accelerometer measurement," in Proceedings of the IEEE International Conference on Mechatronics and Automation, China, 05-08 August 2018 (ICMA IEEE), 1227–1232.
- Hansen, P., Alismail, H., Browning, B., and Rander, P. (2011a). "Stereo visual odometry for pipe mapping," in Proceedings of the IEEE International Conference on Intelligent Robots and Systems, USA, 25-30 September 2011 (IEEE), 4020–4025.
- Hansen, P., Alismail, H., Rander, P., and Browning, B. (2011b). "Monocular visual odometry for robot localization in LNG pipes," in IEEE International Conference on Robotics and Automation, China, 09-13 May 2011 (IEEE), 3111–3116.
- Hansen, P., Alismail, H., Rander, P., and Browning, B. (2013). "Pipe mapping with monocular fisheye imagery," in Proceedings of the IEEE International Conference on Intelligent Robots and Systems, Japan, 03-07 November 2013 (IEEE), 5180–5185.
- Hansen, P., Alismail, H., Rander, P., and Browning, B. (2015). Visual mapping for natural gas pipe inspection. *Int. J. Robotics Res.* 34, 532–558. doi:10.1177/0278364914550133
- Kazeminasab, S., Sadeghi, N., Janfaza, V., Razavi, M., Ziadidegan, S., and Banks, M. K. (2021). Localization, mapping, navigation, and inspection methods in in-pipe robots: A review. *IEEE Access* 9, 162035–162058. doi:10.1109/access.2021.3130233
- Kim, S. H., Lee, S. J., and Kim, S. W. (2018). Weaving laser vision system for navigation of mobile robots in pipeline structures. *IEEE Sensors J.* 18, 2585–2591. doi:10.1109/jsen.2018.2795043
- Kolesnik, M., and Baratoff, G. (2000). Online distance recovery for a sewer inspection robot. *Proc. 15th Int. Conf. Pattern Recognit.* 1, 504–507.
- Lee, D. H., Moon, H., and Choi, H. R. (2016). Landmark detection methods for in-pipe robot traveling in urban gas pipelines. *Robotica* 34, 601–618. doi:10.1017/s0263574714001726
- Ma, K., Schirru, M. M., Zahraee, A. H., Dwyer-Joyce, R., Boxall, J., Dodd, T. J., et al. (2017b). "Robot mapping and localisation in metal water pipes using hydrophone induced vibration and map alignment by dynamic time warping," in Proceedings of the IEEE International Conference on Robotics and Automation, Singapore, 29 May 2017 - 03 June 2017 (IEEE), 2548–2553.
- Ma, K., Schirru, M., Zahraee, A. H., Dwyer-Joyce, R., Boxall, J., Dodd, T. J., et al. (2017a). "PipeSLAM: Simultaneous localisation and mapping in feature sparse water pipes using the rao-blackwellised particle filter," in Proc. of the IEEE/ASME International Conference on Advanced Intelligent, Mechatronics, 03-07 July 2017 (IEEE), 1459–1464.
- Ma, K., Zhu, J., Dodd, T., Collins, R., and Anderson, S. (2015). "Robot mapping and localisation for feature sparse water pipes using voids as landmarks," in *Towards autonomous robotic systems (TAROS 2015), lecture notes in computer science* (Germany: Springer), 9287, 161–166.
- Mateos, L. A., and Vincze, M. (2011). Dewalop-robust pipe joint detection. *Proc. Int. Conf. Image Process. Comput. Vis. Pattern Recognit. (IPC'V)* 47, 1. doi:10.3182/20120905-3-HR-2030.00122
- Matsui, K., Yamashita, A., and Kaneko, T. (2010). "3-D shape measurement of pipe by range finder constructed with omni-directional laser and omni-directional camera," in Proc. of the IEEE International Conference on Robotics and Automation, Anchorage, AK, USA, 03-07 May 2010 (IEEE), 2537–2542.

that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Murtra, A. C., and Mirats Tur, J. M. (2013). "IMU and cable encoder data fusion for in-pipe mobile robot localization," in Proceedings of the IEEE Conference on Technologies for Practical Robot Applications, USA, 22-23 April 2013 (TePRA IEEE).
- Myrans, J., Everson, R., and Kapelan, Z. (2018). Automated detection of faults in sewers using CCTV image sequences. *Automation Constr.* 95, 64–71. doi:10.1016/j.autcon.2018.08.005
- Nassiraei, A. A. F., Kawamura, Y., Ahrary, A., Mikuriya, Y., and Ishii, K. (2006). "A new approach to the sewer pipe inspection: Fully autonomous mobile robot "KANTARO," in Proceedings of the 32nd Annual Conference on IEEE Industrial Electronics, France, 06-10 November 2006 (IECON IEEE), 4088.
- Nguyen, T. L., Blight, A., Pickering, A., Barber, A. R., Boyle, J. H., Richardson, R., et al. (2022). Autonomous control for miniaturized mobile robots in unknown pipe networks. *Front. Robotics AI* 9, 997415–997417. doi:10.3389/frobt.2022.997415
- Pan, X., Ellis, T. J., and Clarke, T. A. (1995). Robust tracking of circular features. *Proc. Br. Mach. Vis. Conf.* 95, 553–562. doi:10.5244/C.9.55
- Powers, D. (2011). Evaluation: From precision, recall and f-measure to ROC, informedness, markedness and correlation. *J. Machine Learning Technologies* 2, 37–63. doi:10.48550/arXiv.2010.16061
- Rizzo, C., Seco, T., Espelósín, J., Lera, F., and Villarroel, J. L. (2021). An alternative approach for robot localization inside pipes using RF spatial fading. *Robotics Aut. Syst.* 136, 103702. doi:10.1016/j.robot.2020.103702
- Roh, S., Lee, J.-S., Moon, H., and Choi, H. R. (2008). "Modularized in-pipe robot capable of selective navigation inside of pipelines," in Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, France, 22-26 September 2008 (IEEE), 1724–1729.
- Rome, E., Hertzberg, J., Kirchner, F., Licht, U., and Christaller, T. (1999). Towards autonomous sewer robots: The MAKRO project. *Urban Water* 1, 57–70. doi:10.1016/s1462-0758(99)00012-6
- Sahli, H., and El-Sheimy, N. (2016). A novel method to enhance pipeline trajectory determination using pipeline junctions. *Sensors* 16, 567–617. doi:10.3390/s16040567
- Schempf, H., Mutschler, E., Gavaert, A., Skoptsov, G., and Crowley, W. (2010). Visual and nondestructive evaluation inspection of live gas mains using the explorer" family of pipe robots. *J. Field Robotics* 27, 217–249. doi:10.1002/rob.20330
- Seco, T., Rizzo, C., Espelósín, J., and Villarroel, J. L. (2016). A robot localization system based on RF fading using particle filters inside pipes. *Proc. Int. Conf. Robot Syst. Compet. (ICARSC)*, 28–34. doi:10.1109/ICARSC.2016.22
- Wang, C., and Huang, K. (2015). How to use bag-of-words model better for image classification. *Image Vis. Comput.* 38, 65–74. doi:10.1016/j.imavis.2014.10.013
- Worley, R., Ma, K., Sailor, G., Schirru, M., Dwyer-Joyce, R., Boxall, J., et al. (2020a). Robot localization in water pipes using acoustic signals and pose graph optimization. *Sensors* 20, 5584. doi:10.3390/s20195584
- Worley, R., Yu, Y., and Anderson, S. (2020b). "Acoustic echo-localization for pipe inspection robots," in Proc. of the IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems, Germany, 14-16 September 2020 (IEEE), 160–165.
- Wu, D., Chatzigeorgiou, D., Youcef-Toumi, K., and Ben-Mansour, R. (2015). Node localization in robotic sensor networks for pipeline inspection. *IEEE Trans. Industrial Inf.* 12, 809–819. doi:10.1109/tii.2015.2469636
- Wu, Y., Mittmann, E., Winston, C., and Youcef-Toumi, K. (2019). "A practical minimalism approach to in-pipe robot localization," in Proceedings of the American Control Conference, USA, 10-12 July 2019 (ACC IEEE), 3180.
- Yu, Y., Worley, R., Anderson, S., and Horoshenkov, K. V. (2023). Microphone array analysis for simultaneous condition detection, localization, and classification in a pipe. *J. Acoust. Soc. Am.* 153, 367–383. doi:10.1121/10.0016856
- Yuen, H., Princen, J., Illingworth, J., and Kittler, J. (1990). Comparative study of Hough transform methods for circle finding. *Image Vis. Comput.* 8, 71–77. doi:10.1016/0262-8856(90)90059-e
- Zhang, R., Evans, M. H., Worley, R., Anderson, S. R., and Mihaylova, L. (2021). "Improving SLAM in pipe networks by leveraging cylindrical regularity," in Proc. Of the annual conference towards autonomous robotic systems (Germany: Springer), 56–65.

Chapter 4

Multiple Hypothesis Particle Filtering For Visual Odometry

4.1 Summary of research

The publication included in the chapter is a short conference article detailing a novel method of navigating a map consisting of junctions and the traversable connections between them. At the time of this research being conducted, the available in-pipe footage was very limited and no method of performing odometry or SLAM within them existed beyond a tether attached to the camera platform which reported how much tether had been un-spooled as an estimate of the distance travelled down the pipe. Therefore, the data used was footage of a car on roads from a standard data-set due to the belief at the time that the topology of road junctions and their connecting roads closely resembled that of pipe networks with manholes connected by pipes and the availability of SLAM and odometry systems able to operate on road networks.

While this belief held mostly true, increased access to pipe network data revealed some further dissimilarities between the two topologies. One of the main differences is the possibility of travelling in both directions along roads and the ability to revisit previously visited junctions. This differs from sewer pipe networks which are designed to allow flow in a single direction along multiple branches towards a single point. This makes pipes directed radial networks as apposed to the undirected mesh network of roads. Despite this, much of the research in the publication remains applicable to pipe networks.

Additionally, pipe networks were found to have additional structure due being constructed

from prefabricated sections of equal length. This structure was found to be detectable and its use for SLAM was explored in the publication contained in chapter 3.

The publication details the use of particle filtering to loosely constrain the vehicle location to the given map and the instantiation of multiple independent particle filters when at a junction. These filters are each assigned to a different possible route taken after the junction and culled as their accuracy deteriorates, leaving the filter that assigned to the correct route to continue. This allows the vehicle to operate past junctions when it has no information about which exit it has taken.

While the primary contribution of this publication is the multi hypothesis particle filtering system, it was also used as an opportunity to implement and trial a deep learning based method of odometry using optical flow. This method was investigated due to it not requiring the feature extraction and matching which causes other methods to fail in pipes as well its purported efficacy and potential to run dedicated optical flow processing hardware designed for mobile robots. The method was ultimately found to be insufficient, however it performs adequately as a stand in for an arbitrary odometry system in any future work.

The publication makes an assumption that a separate system would be capable of identifying when a robot in a pipe network is at a junction (manhole). While no such system had been developed at the time due to a lack of available data, investigation into the structure of pipe environments had led us to be confident in the possibility of such a system. Because of this, and due the increased complexity and diversity among road junctions, it was decided to not attempt such detection during this research. As shown in the previous chapter, this assumption proved to be valid, as an accurate manhole detection system was developed for that publication.

The particle filtering system was found to improve the accuracy of localisation compared to the raw odometry outputs, bringing the mean positional error from 335.2m to 49.2m and the mean heading error from -1.65 radians (-94.54 degrees) to 0.28 radians (16.04 degrees). The system also allows recovery from failure, the lack of which causes the localisation output to be unable to leave an incorrectly entered section of the map when naively integrating prior information.

4.2 Publication

Toward robust visual odometry using prior 2D map information and multiple hypothesis particle filtering*

S. Edwards¹, L. Mihaylova¹, J. M. Aitken¹, and S. Anderson¹

Department of Automatic Control and Systems Engineering, University of Sheffield,
Sheffield, S1 3JD, UK.

Abstract. Visual odometry can be used to estimate the pose of a robot from current and recent video frames. A problem with these methods is that they drift over time due to the accumulation of estimation errors at each time-step. In this short paper we propose and briefly demonstrate the potential benefit of using prior 2D, top-down map information combined with multiple hypothesis particle filtering to correct visual odometry estimates. The results demonstrate a substantial improvement in robustness and accuracy over the sole use of visual odometry.

Keywords: Visual odometry · Deep learning · Multiple Hypothesis · Particle filter · Map prior.

1 Introduction

Visual odometry (VO) is a popular method of pose estimation in mobile robots and there are many methods for this including key-frame optimisation [8] and recently deep learning [6, 3]. The deep learning methods are advantageous because they avoid the need for camera calibration and online optimisation used in key-frame methods, although they tend to be less accurate than the optimisation methods.

One problem that is common to all VO methods (and indeed all odometry methods) is that the pose estimate drifts over time due to the accumulation of estimation errors. Yet, there is often additional information we can use to help reduce drift, such as prior map information. This is true in scenarios of driver-less cars (road maps), mobile robots moving along corridors in indoor environments (architectural floor-plans), and pipe inspection robots such as in the oil, gas, sewer/water and nuclear industries (where pipe network plans tend to be readily available). In all these cases we potentially have prior information in the form of a top-down 2D view of the map, and the mobile agent is largely constrained to move along routes in this map. So, it appears attractive to make

* This work is supported by the UK's Engineering and Physical Sciences Research Council (EPSRC) Programme Grant EP/S016813/1 Pervasive Sensing for Buried Pipes (Pipebots) <https://pipebots.ac.uk>

2 S. Edwards et al.

use of this information in scenarios where it is available. This idea is used in a VO system in [1], where the probability of being located in a discrete road segment is estimated using a particle filter.

The algorithm developed in this paper fuses VO with prior 2D map information using multiple hypothesis particle filtering: when a moving agent reaches a junction in the map, multiple particle filters are used to fuse the VO data with each possible route away from the junction. The most likely hypothesis is probabilistically selected using the distribution of particles for each filter, which acts as a likelihood function, similar to methods that have been used in multiple model particle filtering for fault detection [4]. The results demonstrate that simply using knowledge of the map alone with VO does not lead to successful pose estimation but instead a multiple hypothesis method must be used to ensure accuracy and robustness.

2 Methods

The proposed method of map hypothesis switching is intended for robots or vehicles operating in environments that highly constrain the agent’s motion but whose exact layout is uncertain. Examples of such situations are road or pipe networks where the location of junctions and the connections between junctions are approximately known and available in the form of a 2D, top-down map.

The system performs VO via a deep network similar to those presented in [6, 5] using optic flow calculated via the Horn-Shunck method. The VO outputs are filtered via a particle filter that uses accelerometer and gyroscope data as the basis for a state-space model.

The map provided to the system is represented as a set of coordinates of known junctions with connections between each junction. The system uses the straight lines between each pair of connected junctions as its prior map. Multiple models are instantiated upon reaching a junction and we make one important assumption that the system can identify when it has reached a junction via a separate system. A standard method of multiple model particle filtering, as in [4], is adapted here to switch between particle filter models that express different hypotheses, where the distribution of particles for each filter acts as a likelihood function, and this is used to probabilistically select the most likely hypothesis.

We use the KITTI data set [2] consisting of camera and GPS data from a car, to both train the visual odometry and test the multi hypothesis system. We use different sequences for training and testing, to ensure testing is independent.

3 Results

In order to evaluate our proposed algorithm we tested 1. a system that used VO only, 2. a VO system using a map with a single particle filter, and 3. a VO system using a map with multiple hypothesis particle filtering on separate test data. Fig 1 shows the ground truth and pose estimates for each method overlaid on a street map. As can be seen, the raw VO system performs poorly and a single

Toward robust visual odometry using prior 2D map information 3

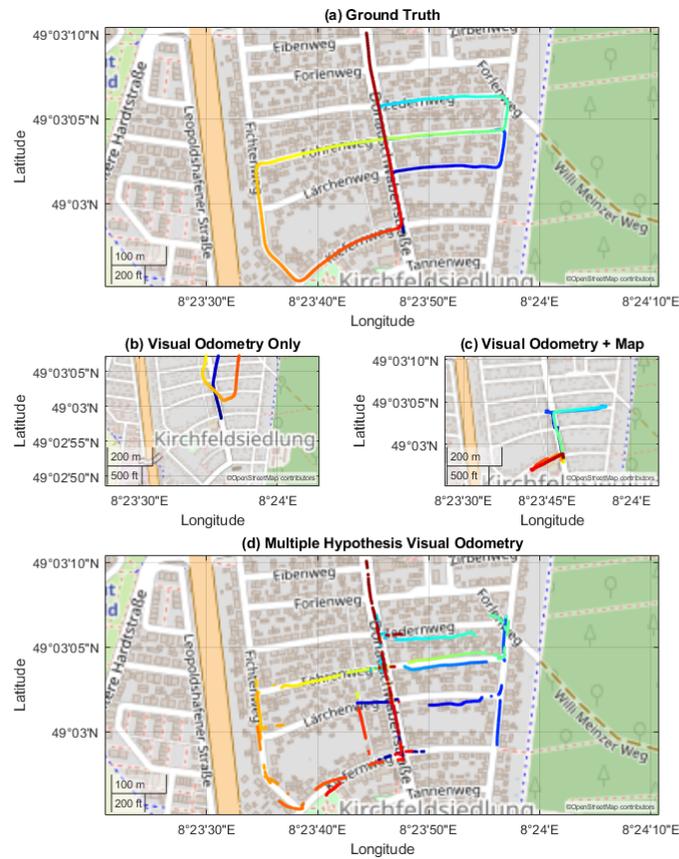


Fig. 1. Comparison of ground truth GPS vehicle data with pose estimation algorithms on independent validation data (KITTI, sequence 5). (a) Ground truth. (b) Visual odometry only (note that the pose estimate exits the area of the ground truth but we retain a comparable zoom-level to the other plots for clarity). (c) Visual odometry with map information. (d) Multiple hypothesis visual odometry.

4 S. Edwards et al.

map hypothesis results in unrecoverable failure when the system believes itself to be in the wrong section of the map. The multiple hypothesis system is generally much more accurate. This system occasionally selects an incorrect hypothesis (corresponding to the apparent gaps in location estimates in Fig. 1d) but then automatically recovers. The principal improvement of this system over simple odometry is the ability to recover from failure, however it also demonstrates an improved pose accuracy, with VO alone resulting in a mean position and heading error of 335.2m(343.20 and -1.65(-2.20) radians respectively and a final positional error of 230.4 m, and the multi hypothesis system resulting in mean errors of 49.2m (40.3) and 0.28 (0) radians and a final error of 0.85 m.

4 Conclusions

In this work we have presented a novel VO method that uses prior 2D map information and multiple hypothesis particle filtering. We demonstrated that the method was more accurate and robust than solely using VO, and that using multiple hypothesis particle filtering substantially improved on using a single particle filter with map information, particularly in its ability to recover from errors. In future work we aim to resolve the problem of junction recognition in order to make a fully standalone algorithm and also develop a compact system that can run in real-time on mobile hardware.

Additionally we aim to address the system's main weakness, that of temporarily selecting an incorrect hypothesis, which occurs regularly after junctions. Improvements here may come from including additional information in the probability calculations, such as designed or learned features from each hypotheses parameters similar to [7], or from more complex assessments of each hypotheses probabilities compared to each other and possibly the system's states and sensor inputs.

References

1. Brubaker, M.A., Geiger, A., Urtasun, R.: Map-based probabilistic visual self-localization. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **38**(4), 652–665 (2016)
2. Geiger, A., Lenz, P., Stiller, C., Urtasun, R.: Vision meets robotics: The KITTI dataset. *The International Journal of Robotics Research* **32**(11), 1231–1237 (2013)
3. Han, L., Lin, Y., Du, G., Lian, S.: DeepVIO: Self-supervised Deep Learning of Monocular Visual Inertial Odometry using 3D Geometric Constraints. In: 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 6906–6913 (2019)
4. Kadiramanathan, V., Li, P., Kirubarajan, T.: Sequential Monte Carlo filtering vs. the IMM estimator for fault detection and isolation in nonlinear systems. In: Willett, P.K., Kirubarajan, T. (eds.) *Component and Systems Diagnostics, Prognosis, and Health Management*. vol. 4389, pp. 263 – 274. International Society for Optics and Photonics, SPIE (2001)

5. Konda, K., Memisevic, R.: Learning visual odometry with a convolutional network. In: The 10th International Conference on Computer Vision Theory and Applications (VISAPP). pp. 486–490 (2015)
6. Muller, P., Savakis, A.: Flowdometry: An optical flow and deep learning based approach to visual odometry. In: 2017 IEEE Winter Conference on Applications of Computer Vision (WACV). pp. 624–631. IEEE (2017)
7. Ossig, D.L., Kurzenberger, K., Speidel, S.A., Henning, K.U., Sawodny, O.: Sensor fault detection using an extended kalman filter and machine learning for a vehicle dynamics controller. In: IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society. pp. 361–366. IEEE (2020)
8. Scaramuzza, D., Fraundorfer, F.: Visual Odometry [Tutorial]. IEEE Robotics & Automation Magazine **18**(4), 80–92 (2011)

Chapter 5

Conclusions

This thesis aimed to address the problem of autonomous navigation in buried pipe networks by leveraging prior knowledge of pipe environments. Section 1.2 stated five objectives to met in order to realise this aim. The research presented in the previous chapters has met those objectives, and in doing so, made significant contributions to the field of SLAM in pipe networks.

The first of those objectives (the creation of usable maps of pipe networks) has been addressed in chapters 3 and 4 via the creation of maps used in the corresponding publications. Chapter 4 does this by approximating road maps to pipe networks while chapter 3 operates in real pipes and uses knowledge of their structure to create a map suitable for its method of odometry.

The system of odometry via landmark detection detailed in chapter 3 effectively addresses the second and objective of this thesis, presenting systems of detecting landmarks and manholes with 85% and 98.5% accuracy's respectively. It is then able to use those detections to locate itself within its environment relative to its last known location and obtain the path it has taken with an average error of 1.76m, thus addressing the third and fourth objectives significantly effectively than ORB-SLAM3's average error of 11.53m in the same environment.

Chapter 4 then completes the final objective by developing a method of robustly considering multiple possible paths in order to account for uncertainty in the direction travelled and to allow for odometry systems to recover from failure. This method improves the average error of the odometry system alone from 335.2m to 49.2m.

While autonomous navigation of pipe networks remains a complex and challenging problem, this thesis presents promising progress toward an overall solution. The system described in chapter 3 is able to localise within a pipe network to a degree that, to this authors knowledge,

has not previously been demonstrated. This system could be used in real world navigation systems with only minor optimisation of its parameters for the intended operating environment. It could also be further developed to include other predictable landmarks such as inlets, or be adapted for other types of pipes provided they have suitable structural landmarks. The work in chapter 4 requires adapting specifically to pipe networks, however it serves as a foundation for building such a system and demonstrates the effectiveness of map integration by multiple hypothesis particle filtering.

A continuation of this thesis would likely begin by applying the multi hypothesis particle filtering algorithm developed in chapter 4 to pipe network odometry. While the existing method should, in theory, work equally well in a pipe network, some work may be needed to adapt and tune the algorithm for this environment. Work would also need to be conducted to integrate the algorithms developed in chapters 3 and 4 into a single system.

The most obvious continuation however, would likely be to address the problem of creating maps of pipe networks since the research included in this thesis only used existing maps or knowledge of the environment. Although some work on this topic was undertaken, it was not up to publishable standards and therefore this thesis did not present research on the problem of creating maps of pipe networks from above ground detections of manholes. It is the authors opinion that low altitude aerial imagery is a potential solution to this problem however more research would be needed to assess the viability of various detection methods as well as methods of converting those detections into a usable map.