

FEATURE DETECTION AND CLASSIFICATION IN BURIED PIPES USING LiDAR TECHNOLOGY

Aristeidis Karnezis¹, Rob Worley², Andy Blight³, Sean Anderson⁴, Kirill Horoshenkov⁵, Lyudmila Mihaylova⁶

^{1,2,4,6} School of Electrical and Electronic Engineering, University of Sheffield

³ School of Mechanical Engineering, University of Leeds

⁵ School of Mechanical, Aerospace and Civil Engineering, University of Sheffield

[¹a.karnezis@sheffield.ac.uk](mailto:a.karnezis@sheffield.ac.uk)

ABSTRACT

Buried infrastructure presents unique challenges for autonomous robotic inspection due to its confined geometry and the structural variations within pipe networks. While CCTV is widely used for pipe inspection, LiDAR sensors offer complementary advantages, including precise ranging capabilities and accurate depth perception. In this work, we introduce a low-cost LiDAR-based system designed to detect blockages and accurately identify critical structural features - such as joints, manholes, and other discontinuities - within these environments in real-time. By combining robust data acquisition, efficient processing, and clear decision-making criteria, the approach enhances the effectiveness, reliability, and automation of underground inspections.

Keywords: LiDAR, Feature detection, Decision making

INTRODUCTION

Buried sewer and wastewater pipes form an essential part of city infrastructure, with 0.8 million kilometers in the UK. Inspection is essential but costly, and is predominantly done by deploying tethered crawlers through manholes which are remotely controlled by an operator on the roadside. The structure and condition of the pipes are assessed using video feeds recorded by the crawler [1]. Introducing a LiDAR scanner, which uses a laser to measure the distance from the sensor to surrounding surfaces, would enhance defect detection and enable autonomous navigation in future untethered robotic systems.

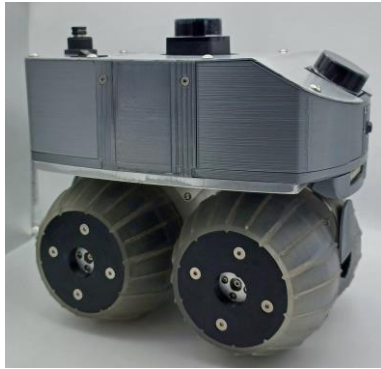
The pipe environment imposes a number of constraints on inspection robots. The confined pipe geometry limits space for sensing, so computation should be applicable to small devices. The possible lack of available communications means that any computation for low-level control must be done on-board the robot. Lighting conditions are also minimal, making vision-based methods unreliable without additional illumination. For untethered, long-duration missions, both sensing and computation must be power-efficient to conserve battery life. At the same time, there are some aspects of the application which offer advantages to LiDAR-based sensing. The robot motion can be slow and steady, especially during manoeuvring. Therefore, the rate of sensing and localization required is low compared to aerial or open-environment robots. The environment is expected to be relatively static comprising only a small range of possible local environments: principally pipes and manholes, which is useful for classification [2].

In our application, pipes are around 300mm in diameter, and will generally be tens of metres long. Pipes are connected at manholes, which are larger chambers, greater than 1 metre across. Each of these possible environments

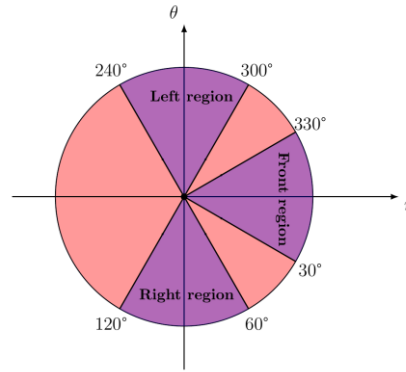
can be relatively featureless and self-similar, so the information that can be obtained from them is limited. The features that do exist are joints between sections of pipe, which are typically spaced around 3 metres apart, and blockages, which will vary in size and position.

METHODS

Our system is built around the low-cost Okdo LiDAR Module (LD06) [3], mounted at a 30° angle relative to the horizontal axis. This provides a good balance between detection range and point density making detection reliable and computationally efficient. As shown in Figure 1(a), the sensor is integrated into our robotic platform and positioned 150mm above ground level, allowing us to effectively capture data from approximately 256mm ahead of the robot. Although the LiDAR captures a full circular scan, we only process data within the three 60° regions shown in Figure 1(b), where obstacles and structural features are most likely to appear. The LiDAR emits laser pulses inside a pipe, recording the time-of-flight, angle, range and intensity of each reflection.



(a)



(b)

Figure 1. Robotic platform with the LiDAR sensor mounted at a 30 degrees angle and illustrates how each LiDAR scan is divided into three regions - front, left and right - for feature and blockage detection.

Data Preprocessing. To enhance data quality, we first filtered raw LiDAR measurements, discarding invalid points and entries with infinite values. The remaining filtered points, each consisting of a radial distance r_i and a polar angle θ_i are projected into the horizontal plane using our known sensor tilt. These projections are converted to Cartesian coordinates to give corresponding (x_i, y_i) pairs and grouped to the three regions of interest shown in Figure 1(b).

Fast Two-Stage Decision System using Thresholding. The LiDAR-based feature detection needs to operate in real-time on a low-powered robot, therefore, we develop a lightweight, fast decision algorithm.

Stage 1 Clustering. Within each region of interest (front, left, right), we apply the DBSCAN algorithm to identify k spatially coherent clusters of points C_k along that slice of the pipe [4]. For each cluster C_k , we compute the average horizontal distance d_k from the robot to the LiDAR reflective points in that cluster

$$d_k = \frac{1}{|C_k|} \sum_{(x_i, y_i) \in C_k} \sqrt{x_i^2 + y_i^2},$$

where (x_i, y_i) are the cluster's projected coordinates. These mean distances are then sorted in ascending order ($d_1 \leq d_2 \leq \dots$). The cluster corresponding to the smallest mean distance d_1 is chosen as our primary candidate and by extension the nearest surface. This cluster is accepted only if the separation distance $\delta = d_2 - d_1$ (gap between the nearest and the next-closest cluster) exceeds a predefined threshold, otherwise, the cluster is discarded.

Stage 2 Decision Threshold. Once the nearest cluster has been selected via its smallest average horizontal distance d_1 and cleared the gap test against the next-closest cluster, we apply a threshold-based classifier to distinguish pipe walls, joints, manholes, and when looking straight ahead of the robot, blockages. In the right and left regions, the closest clusters are first assigned to pipe-wall when they lie within the typical pipe radius. Clusters that fall just beyond that range are interpreted as joints (the small gap at the interface of two pipe segments), and anything still farther out are treated as manhole openings. Any measurement that does not clearly fall into one of these ranges is treated as ambiguous and discarded. In the front region, once a cluster passes the gap test, we take the smallest average horizontal distance d_1 and compare it to an obstruction threshold. If d_1 falls below that threshold, we classify the cluster as a blockage and immediately alert the control system, otherwise, no front detection is issued, indicating a clear path ahead.

LiDAR Ground Truth Labelling. To generate reliable ground truths, we developed a semi-automated labelling system. Raw LiDAR scans were visualised in polar form, and structural features were manually annotated. Each annotation captures the scan identifier, radial distance r_i and polar angle θ_i of the LiDAR beam as well as the label of each feature. To validate generality, data were drawn from a variety of pipe materials, resulting in a balanced set of several hundred labelled features. From this dataset, we formulated a simple binary classification problem: distinguishing joints from manholes using each scan's (r_i, θ_i) as input features. This high-quality reference dataset allowed us to evaluate the performance of our Fast Two-Stage Decision System and also train and benchmark against a Random Forest classifier [5], discussed below.

Performance Evaluation. To establish a baseline accuracy on feature classification performance using a more computationally intensive classifier, we trained a Random Forest ensemble composed of 200 decision trees with balanced class weights on 80% of our manually annotated scans (split so that no scan appeared in both sets). We then evaluated the model on the remaining 20% and generated a normalised confusion matrix to quantify its classification accuracy for a subset of the four classes: joints and manholes. We also ran the Fast Two-Stage Decision System on the same LiDAR data. For each region in every scan, the detector generated a radial distance d , a polar angle θ and a predicted label. Each prediction (d, θ) was matched to the nearest ground-truth annotation (r_i, θ_i) and considered correct if $|d - r_i| \leq \epsilon_d$ and $|\theta - \theta_i| \leq \epsilon_\theta$ where ϵ_d and ϵ_θ are the predefined distance and angle tolerances.

RESULTS AND DISCUSSION

Having outlined our sensor setup, clustering method, and two-stage decision system, we then evaluated how these components perform in practice.

Figure 2 illustrates a 3D spatial reconstruction of an 11.5m pipe network with the corresponding temporal outcomes of the clustering method. In the top panel LiDAR scans are projected into the horizontal plane and aligned using the robot's wheel-odometry to build a continuous 3D model of the pipe interior.

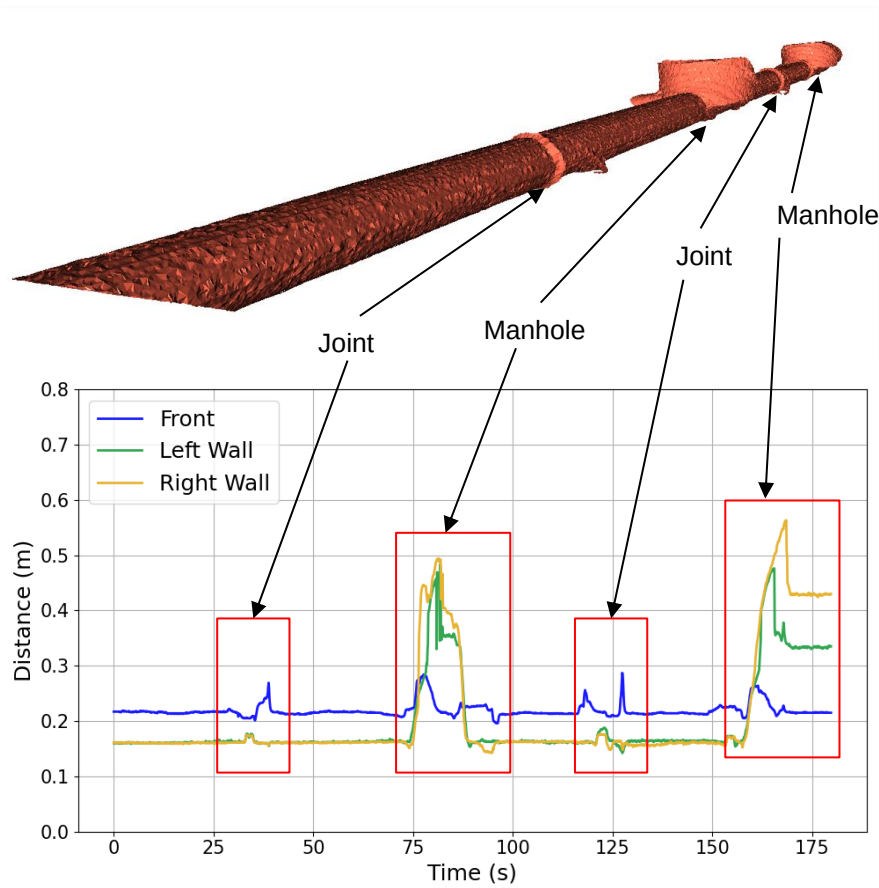


Figure 2. 3D reconstruction of a pipe network using LiDAR scans and robot position estimates, alongside a time series graph of the mean horizontal distance from the front, left- and right-wall regions to the surface of the pipe.

In this reconstruction, pipe joints between pipe segments appear as thin circular bands, while manhole chambers are visible as distinct expansions in the pipe's cross-section. As the robot travels through the pipe, each LiDAR scan is split into front, left-wall, and right-wall clusters. In the bottom panel, the time series of the smallest average horizontal distance in the front (blue), left-wall (green), and right-wall (orange) regions remains flat through straight runs and spikes at joint and manhole locations, demonstrating real-time detection of structural features.

Next, we examined the side-wall distance distributions to validate our threshold values. Figure 3(a) presents a histogram of the average horizontal distances from the left-wall and right-wall regions, overlaid with two fitted Gaussian curves: one centered at the true pipe radius, representing pipe-wall scans and a second centered at the joint distance, representing joint scans. The clear separation between these peaks confirms that our predefined distance threshold reliably distinguishes between the two feature classes. Figure 3(b) presents the histogram of manhole distance measurements with two separate Gaussian fits. These fits differ because the LiDAR is tilted within the manhole and thus its horizontal-plane slice cuts the circular opening above its centre,

producing two unequal chord lengths. Both manhole distributions lie well above the joint and pipe-wall thresholds, confirming that a higher distance cutoff cleanly isolates manholes.

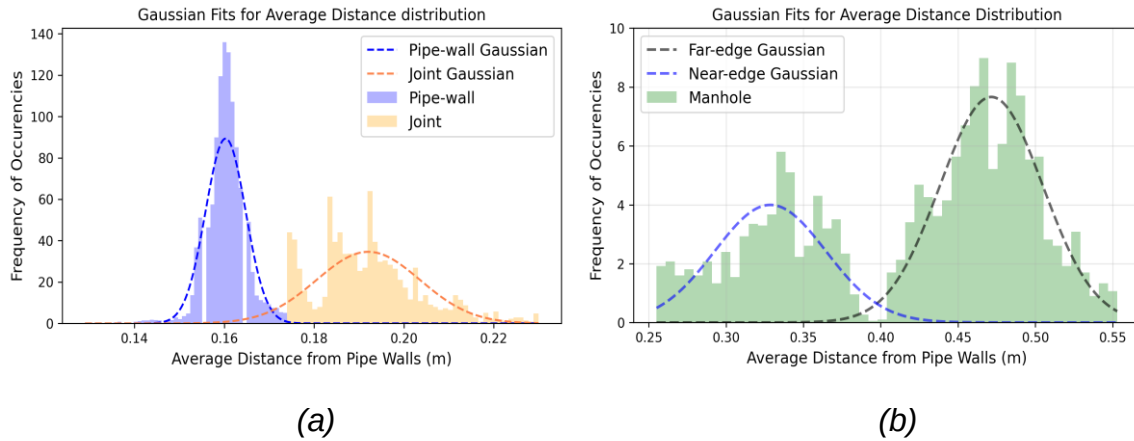


Figure 3. (a) Histogram of side-wall average distances with Gaussian fits for the pipe-wall and joint clusters. (b) Histogram of manhole side-wall distances with two Gaussian fits for the shorter and longer chord lengths.

Having confirmed that our distance threshold separated pipe walls, joints and manholes, we then measured how well those thresholds and our Fast Two-Stage Decision System performed in a real-pipe environment. We assessed the binary classification performance on the two main feature classes - manhole and joint. The Random Forest classifier achieved 97% recall for manholes and over 93% recall for joints, as shown in Figure 4(a), confirming that a flexible, data-driven model can reproduce the distinctions between joints and manholes. Next, we evaluated the performance of our fast detection algorithm by matching its outputs directly against the ground truth annotations. Each predicted label was paired with the nearest ground truth annotation within $\pm 5mm$ in radial distance and $\pm 1^\circ$ in polar angle allowing each prediction to be uniquely matched to its annotation. The resulting confusion matrix, presented in Figure 4(b), demonstrates recall rates of 96.9% for manholes and 95.5% for joints, indicating that the combination of clustering and an efficient threshold-driven classifier provides robust real-time detection and accurate classification of pipe features in confined pipe environments.

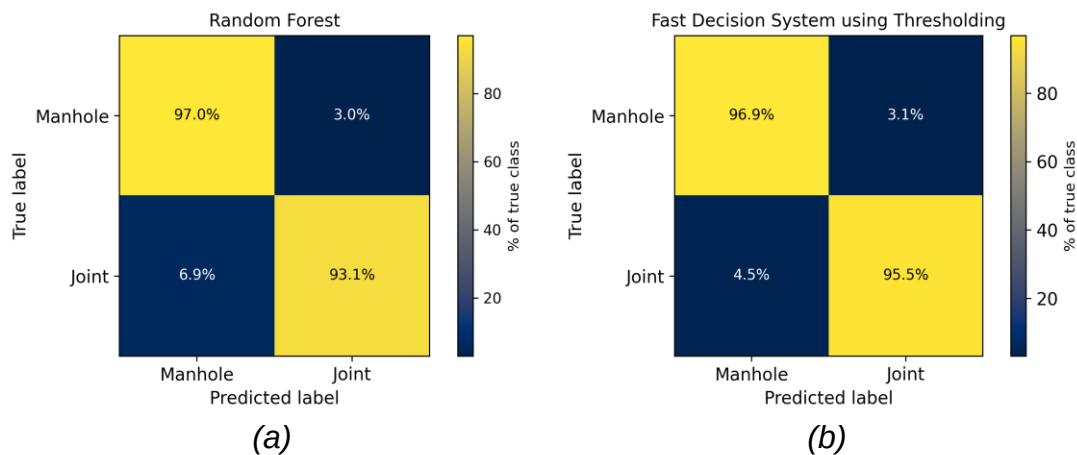


Figure 4. Confusion Matrices for manhole and joint classification from: (a) Random Forest trained on ground truth annotations. (b) The two-step feature detection algorithm.

CONCLUSIONS

This paper has presented a low-cost LiDAR-based approach for real-time detection and classification of critical structural features within buried pipe networks. By combining robust pre-processing, spatial clustering using DBSCAN, and a clear two-step decision framework, our detection algorithm reliably distinguishes between smooth pipe walls, joints and manholes, matching the performance of a supervised Random Forest benchmark. Future work will expand the binary classification to robustly distinguish between an unobstructed pipe and a blockage before extending the framework to a full multiclass classifier for a variety of feature types.

ACKNOWLEDGEMENTS

We are grateful for the support via the EU project: “PIPEON: Robotics and AI for Sewer Pipe Inspection and Maintenance” under grant number 101189847.

REFERENCES

- [1] M. Ebrahimi, H. Hojat Jalali, and S. Sabatino, ‘Probabilistic condition assessment of reinforced concrete sanitary sewer pipelines using LiDAR inspection data’, *Automation in Construction*, vol. 150, p. 104857, Jun. 2023.
- [2] F. Kirchner and J. Hertzberg, ‘A Prototype Study of an Autonomous Robot Platform for Sewerage System Maintenance’, *Auton. Robots*, vol. 4, no. 4, pp. 319–331, Oct. 1997.
- [3] Okdo LiDAR Module LD06, Datasheet available [online](#), Visited 6 July, 2025.
- [4] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, ‘A density-based algorithm for discovering clusters in large spatial databases with noise’, in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, in KDD’96. Portland, Oregon: AAAI Press, 1996, pp. 226–231.
- [5] J. Myrans, Z. Kapelan, and R. Everson, ‘Automated Detection of Faults in Wastewater Pipes from CCTV Footage by Using Random Forests’, *Procedia Engineering*, vol. 154, pp. 36–41, 2016.