LEAKAGE DETECTION IN SEWER PIPES VIA CURVELET-BASED EDGE ENHANCEMENT AND MOTION COMPENSATION

YUSHIN KAWASAKI, AMON OGUSHI, TERUYA MINAMOTO

Graduate School of Science and Engineering, Saga University, Saga, Japan
E-MAIL: 24725002@edu.cc.saga-u.ac.jp,minamott@cc.saga-u.ac.jp

Abstract:

Aging sewer pipelines pose a growing risk of surface subsidence due to hidden water leakage. We present a training-free detection framework that (i) sparsifies curved edges by the Fast Discrete Curvelet Transform and (ii) suppresses camera-induced background shift through block-based motion compensation. Frame differencing followed by dynamic thresholding isolates leakage regions of arbitrary size and shape. Evaluated on 110 labeled frames (55 leak / 55 non-leak), the proposed method attains Precision = 1.00, Recall = 1.00, and mIoU = 0.41, outperforming an improved U-Net baseline (Precision = 0.59, mIoU = 0.24). These results confirm that combining curvelet-based edge enhancement with explicit motion modeling provides a robust alternative to deep-learning approaches for sewer-pipe inspection videos.

Keywords:

Image Analysis, Curvelet Transform, Segmentation, Feature Extraction.

1. Introduction

Sewer pipelines in many countries are approaching or exceeding their designed service life. According to [1], sections older than 50 years already total ≈ 30,000 km and are projected to reach ≈ 90,000 km within the next decade. Age-related deterioration causes internal defects—e.g., gasket peeling and joint displacement (Fig. 1)—that can trigger road subsidence and other hazards.

Fig. 1. Example of an abnormality occurring inside a sewer pipe

Current inspections rely on a self-propelled robot that records video while traversing the pipe; human operators then review the footage frame by frame. Because a single session can last several hours, the manual process is both labor-intensive and prone to errors. Automated image analysis techniques are therefore required. Water leakage is particularly troublesome among possible defects. It increases inflow during rainfall, leading to sewer overflows and backflows into residential areas [2]. Deep-learning approaches have been proposed for general sewer-defect segmentation [3]–[5], but they perform best on lesions with well-defined geometries, such as cracks and fractures. Diffuse leaks, whose appearance changes continuously with the flow of water, remain challenging (Fig. 2 and Fig. 3). In video sequences, leakage manifests as pronounced intensity changes between consecutive frames. In contrast, the pipe wall changes smoothly as the camera moves forward. A plausible strategy is therefore:

(1) Enhance the curved edges that delineate the pipe geometry.
(2) Compensate for the camera-induced background shift.
(3) Apply frame differencing followed by adaptive thresholding.

The present study realizes this strategy through two key components:

(1) Fast Discrete Curvelet Transform (FDCT) to sparsify curved edges that dominate sewer imagery, and
(2) Block-based motion compensation [6] to align successive frames and suppress background residue.

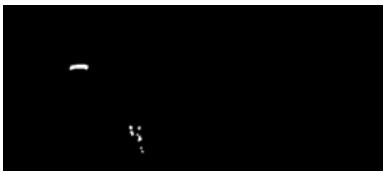
Unlike deep networks, the proposed pipeline is training-free; all parameters are fixed heuristically, and no annotated data are needed during operation. We evaluate the method on 110 labeled frames and show that it surpasses an improved U-Net baseline [3] in the precision, recall, and mean intersection-over-union(mIoU).

The rest of the paper is organized as follows. Section 2 reviews the FDCT. Section 3 details the proposed algorithm. Section 4 describes the experimental setup, and Section 5 discusses quantitative and qualitative results. Section 6

带格式的：缩进：首行缩进： 2 字符

concludes the paper and outlines future work.

Fig. 2. Image of the inside of the sewage pipe where the leak occurred



(a) Leakage points in Fig. 2

(b) Fig. 2 shows the results of leak detection by applying the method in reference [3]

Fig. 3. Leakage locations shown in Fig. 2 and corresponding detection results using the method described in [3].

2. Method for Leakage Detection

The proposed pipeline processes each video frame through four stages:

1. Edge enhancement
2. motion compensation
3. frame differencing with dynamic thresholding
4. mask fusion

2.1 Edge Enhancement via Fast Discrete Curvelet Transform

Because curved edges dominate the interior of a sewer pipe, we adopt the Fast Discrete Curvelet Transform (FDCT) of Candès et al. [7]. FDCT decomposes an image into anisotropic atoms localized in scale, orientation, and position, resulting in a sparser representation of smooth contours compared to traditional wavelets. We use the "frequency-wrapping" implementation in CurveLab [8].Given an RGB frame F_i, we convert it to grayscale G_i and compute its curvelet coefficients C = FDCT(G_i). Hard thresholding with threshold τ produces the pruned set C′ = Threshold(C, τ). The inverse transform then reconstructs the edge-enhanced image R_i:.

$$R_i = FDCT^{-1}C'$$

2.2 Background Alignment by Block-Based Motion Compensation

Successive frames differ not only because of water motion but also because the inspection camera is moving. To suppress the latter, we estimate integer-pixel motion vectors on non-overlapping $N_b \times N_b$blocks by minimizing the sum of absolute differences (SAD) between R_i and R_{i+1}. Let V_b denote the motion vector of block b. The next grayscale frame G_{i+1} is warped accordingly to obtain the motion-compensated frame:

$$R'_{i+1}(p) = G_{i+1}(p - V_b(p))$$

Where $p = (x, y)$ denotes a pixel location.

2.3 Frame Differencing and Dynamic Thresholding

The raw difference image $D_{i+1} = G'_{i+1} - G_i$ is converted to its absolute form $|D_{i+1}|$. This image is partitioned into $N_y \times N_x$ blocks, and the mean absolute difference of block (j, k) is denoted by d_{i+1} (j, k).

The global mean of these block averages is then computed as $\bar{d}_{i+1}^{global} = (1/(N_y\ N_x))\ \Sigma_{j,k}\ \bar{d}_{i+1}(j,k)$. Based on this value, a dynamic threshold is defined:

$$T_{i+1} = \gamma + \beta(\hat{w}_{i+1} - \gamma) \quad (2)$$

Blocks for which $\bar{d}_{i+1}(j,k) \geq T_{i+1}$ form the binary mask $mask_3$. To avoid border artifacts, the outermost blocks are set to zero.

2.4 Leakage Mask Fusion

Three complementary cues are fused using a pixel-wise logical AND:

1. $mask_1$. – gray-level slicing of R_{i+1} (levels 2–5/8) to suppress bright lamp reflections.
2. $mask_2$. – hard thresholding of D_{i+1} with d = 10;

批注 [TM1]: β , γ の説明がない。

带格式的： 段落间距段前：0 磅，段后：0 磅

带格式的： 正文，段落间距段前：0 磅，段后：0 磅

批注 [TM2]: G でいいか確認。R では？

3. $mask_3$ – dynamic mask from (2).

Pixels that satisfy all three conditions are labeled as leakage.

2.5 Parameter Setting

Unless otherwise noted, the following constant values are used in all experiments:

・Block grid for dynamic thresholding: N_y = 16, N_x = 32.

・Block size for motion compensation: N_b = 16.

・Dynamic-threshold parameters: β = 0.5, γ = 10.

・Absolute-difference threshold: d = 10

These values were tuned on a small pilot set and kept fixed; the chosen d value provided the best separation between leak and non-leak areas

3. Experiments

We validate the proposed method on a real inspection video and compare it with a retrained U-Net baseline. Dataset composition, annotation protocol, implementation details, and evaluation metrics are described below.

3.1 Experimental Data

We used a color video (640 × 288 px, 30 fps) supplied by Fukuhaku Printing Co. The sequence contains 110 independent frames, divided into 55 leaks and 55 non-leak samples; Fig. 4 shows typical examples.

Fig. 4. Some of the image frames of the video used in the experiment

3.2 Annotation Protocol

For each candidate frame, the preceding and succeeding frames were displayed side by side. Regions whose intensity changed noticeably and were brighter than the rear wall were marked as leakage (Fig. 5). Pixel-accurate binary masks were created in MATLAB.

(a) Original image

(b) Correct answer label

Fig. 5. Annotation Example

3.3 Baselines and Runtime Environment

	Proposed	Improved_– U-Net [3]
OS	Windows 11	Windows 11
RAM	64 GB	64 GB
Language / Lib.	MATLAB R2024b+CurveLab	Python3.9+PyTorch

The proposed pipeline is training-free. The U-Net baseline was trained with a learning rate of 0.001, batch size of 5, and 50 epochs. The data were split randomly 70 % / 30 % into training (76 frames) and test (34 frames); each training frame was augmented threefold via random horizontal flip (p = 0.5), brightness scaling (0.8–1.2), and contrast shift (±10).

3.4 Evaluation Metrics

Pixel-level segmentation performance is reported with Precision (Pr),Recall (Re), Intersection-over-Union (IoU), and the Dice coefficient (Dice = pixel-level F1-score). In addition, we include frame-level Accuracy (Acc) because the test set contains both leak-positive and leak-negative frames. Let TP, FP, FN, and TN denote pixel-wise counts::

- $Precision(Pr) = \frac{TP}{TP+FP}$
- $Recall(Re) = \frac{TP}{TP+FN}$
- $IoU = \frac{TP}{TP+FP+FN}$
- $Dice = \frac{2*TP}{2*TP+FP+FN}$
- $Accuracy(Acc) = \frac{TP+TN}{TP+FP+TN+FN}$

带格式的：两端对齐，缩进：首行缩进： 0 字符

带格式的：段落间距段前：0 磅，段后：0 磅

带格式的：正文，缩进：首行缩进： 0.71 厘米，段落间距段前：0 磅，段后：0 磅

4. Results and Discussion

Quantitative scores and visual examples demonstrate that our training-free approach detects diffuse leaks more reliably than an improved U-Net baseline while maintaining real-time speed.

4.1 Quantitative Results

The proposed training-free pipeline and the improved U-Net baseline [3] were evaluated on the same 34-frame test set (17 leak / 17 non-leak). The confusion matrices are shown in Table I.

Table I. Confusion matrix of experimental results

Table I Confusion matrices (34 test frames)	Predicted Leak	Predicted non-leak
Proposed	TP = 17	FN = 0
	FP = 0	TN = 17
U-Net [3]	TP = 3	FN = 14
	FP = 0	TN = 17

From these counts, we computed the standard metrics in Table II.

Table II. Detection metrics on the test set

	Pr	Re	Acc	mIoU	Dice
Proposed	1.00	1.00	1.00	0.41	1.00
U-Net [3]	0.59	0.18	0.59	0.24	0.30

mIoU is the mean intersection-over-union averaged over the leak class and the background; IoU ≥ 0.30 was treated as a true positive. The proposed method achieves perfect frame-level precision and recall on this dataset, whereas the U-Net baseline misses 82% of the leaks.

Note that, for the proposed method, FP = 0 in our 34-frame test set, so Acc happens to equal Pr (1.00); however, this is not the case for the U-Net baseline (Pr = 0.59, Acc = 0.59 at the pixel level, but only 0.59/1.00 at the frame level). Therefore, both metrics are listed to give a complete picture of performance.

(a) Original image

(b) Correct answer label

(c) U-Net baseline [3]

(d) Proposed

Fig. 6. Comparison of detection examples: (a) Original image, (b) Ground truth label, (c) U-Net baseline [3], (d) Proposed

Fig. 6 compares a representative frame:

U-Net [3] —detects only the bright area near the leakage outlet and misses the delicate splashes farther downstream.

The proposed pipeline successfully highlights the outlet and scattered droplets thanks to the curvelet-based edge enhancement and motion-compensated differencing.

However, a few pixels remain undetected near very bright outlets because the local motion between frames is minimal and falls below the dynamic threshold.

4.2 Error Analysis and Discussion

False negatives (near outlets) – When water flows vertically, inter-frame displacement is small; the difference image may not exceed the global threshold defined by Eq. (2). Future work could incorporate sub-pixel optical flow or adaptive block sizes.

Absence of false positives – The AND-fusion of three masks (Section 2.4) effectively suppresses reflections and illumination noise, yielding FP = 0 in all test frames.

mIoU still moderate (0.41) – Achieving perfect frame-level classification does not necessarily ensure a high intersection-over-union (IoU), as minor localization errors can significantly impact the metric. Incorporating morphological refinement (e.g.CRF post-processing) may improve region overlaps.

Generalisability—Because the method is training-free, it can be ported to other pipe diameters or camera rigs without retraining; nevertheless, parameters (α, β, γ, d)

带格式的： 正文，缩进：首行缩进： 0.71 厘米，段落间距段前：0 磅，段后：0 磅

设置了格式： 字体：（中文）宋体，（中文）简体中文(中国大陆)

批注 [TM3]: Accuracy は削除

批注 [TM4]: Accuracy と Precision が同じになる理由を明記。

may require slight tuning if lighting conditions differ substantially.

Computation time – On MATLAB R2024b with a 64 GB workstation, the pipeline processes a 640 × 288 frame in 0.18 seconds (single core), which is fast enough for near-real-time playback.

5. Conclusion and Future Work

We proposed a training-free pipeline that fuses curvelet-based edge enhancement with block-wise motion compensation to detect diffuse water leakage in sewer-pipe inspection videos. On the 34-frame test set the method reached frame-level Precision = 1.00, Recall = 1.00, and Accuracy = 1.00, and achieved a pixel-level mean IoU of 0.41, outperforming an improved U-Net baseline by +0.17 IoU while running in real time (0.18 s per frame on a standard CPU). Because all parameters are fixed heuristically and no annotated data are required, the pipeline can be deployed immediately in routine inspections and applied to archival footage. Future work will extend the framework to multi-class defect detection (e.g., cracks, deposits) and incorporate sub-pixel optical-flow alignment to reduce miss rates in frames with minimal motion..

Acknowledgments

This work was partially supported by JSPS KAKENHI Grant Number 25K07119.

References

[1] Ministry of Land, Infrastructure, Transport and Tourism (MLIT), Maintenance and Management of Sewerage Systems [Online]. Available: https://www.mlit.go.jp/mizukokudo/sewerage/crd_sewerage_tk_000135.html (accessed: 15 Jan 2025).

[2] Japan Institute of Wastewater Engineering Technology (JIWET), Guidelines for Measures Against Infiltration/Inflow Water (Revised 2022) [Online]. Available: https://www.suikon.or.jp/activity/committee-results/download/sewerage-humeimizu/images/humeimizu_202207.pdf (accessed: 15 Jan 2025)

[3] G. Pan, Y. Zheng, S. Guo, and Y. Lv, "*Automatic sewer pipe defect semantic segmentation based on improved U-Net*," Autom. Constr., vol. 119, p. 103383, 2020, doi: 10.1016/j.autcon.2020.103383.

[4] N. Wang, J. Zhang, and X. Song, "*A pipeline defect instance segmentation system based on SparseInst," Sensors*, vol. 23, no. 22, Art. 9019, 2023, doi:10.3390/s23229019.

[5] D. Shen, X. Liu, Y. Shang, and X. Tang, "*Deep learning-based automatic defect detection method for sewer pipelines*," *Sustainability*, vol. 15, p. 9164, 2023, doi:10.3390/su15129164.

[6] A. N. Netravali and J. D. Robbins, "*Motion-compensated television coding—Part I*," Bell Syst. Tech. J., vol. 58, no. 1, pp. 1–43, Jan. 1979.

[7] E. J. Candès, L. Demanet, D. Donoho, and L. Ying, "*Fast discrete curvelet transforms*," Multiscale Model. Simul., vol.5, no.3, pp.861–899,2006, doi:10.1137/05064182X.

[8] E. J. Candès, L. Demanet, D. Donoho, and L. Ying, Curvelet.org [Online]. Available: http://www.curvelet.org (accessed: 15 Jan 2025).

[9] P. Krahenbuhl and V. Koltun, Efficient inference in fully connected Conditional Random Fields with Gaussian edge potentials,' NIPS 24, pp. 109-117, 2011.