

Image processing-based framework for determining of growth in sewer pipe  
defects

By

Mohamed Karabij

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Construction Engineering and Management

Department of Civil and Environmental Engineering

University of Alberta

## **Abstract**

Municipal drainage systems have a key role in public health and are considered one of the main components of every modern city's infrastructure. However, as the drainage system ages, its pipes gradually deteriorate at rates that vary based on the different conditions of utilization. To prevent unexpected failures, municipalities have adopted a proactive approach that relies on regular condition assessments of their assets from which data-driven maintenance, rehabilitation, and replacement plans are developed. From a practical standpoint, this data-driven planning relies on assessment information used in conjunction with deterioration models to evaluate the risk of failure associated with structural and operational anomalies. In this respect, data-driven plans can be very useful in assisting municipalities insofar as budget and process management are concerned. The most popular deterioration models rely on statistical analysis and statistical models whose accuracy depends heavily on the quality and the quantity of data collected as part of underground pipe inspections, e.g. the number of cracks, fracture, roots, etc. However, understanding the development of defects over time will not only help improve the accuracy of the existing deterioration models but will also provide valuable information in terms of understanding the relationship between the various factors affecting defect development and pipe deterioration. This research presents an image registration framework for extracting crack development information from the CCTV videos of sewer pipes. Image processing techniques are used to estimate a relative change in a defect from images taken at two different times to determine the relative growth of the defect. The framework was implemented on two case studies and a visual validation was applied to the results. In this respect, since the camera is not expected to be identical for the image time series, the accuracy of Area Scaler (AS), which is used so as to have a similar scale between the

images, is examined. This scaling procedure is illustrated using a case study, containing 49 pairs of images, leading to a 95.5% accuracy.

## **Acknowledgments**

I wish to express my gratitude to my supervisors Dr. Ahmad Bouferguene and Dr. Mohamed Al-Hussein for their continuous encouragement, support, and valuable advice throughout all the stages of this research. I also would like to express my sincere appreciation to the many wonderful people who have supported me during my studies, both directly and indirectly, through friendship and inspiration, including all my colleagues in EPCOR project group and the modular construction research group at the University of Alberta.

I owe my deepest gratitude to my parents, who did and are doing the impossible to support my education and push me to pursue my dreams as well as instilling in me the desire to learn and the motivation to succeed.

I am thankful for the help and encouragement of my siblings, Taher, Ubaida, and Karam, during my study.

I would like to thank my best friend in this life, my wife, Marwa Almohamad, for her unconditional caring, encouragement, and support.

I would like to extend my thanks to all the good people who supported me and were a big reason to achieve this work, naming a few, Eng. Ahmed Alrifai, Dr. Hussein Merza, and Mr. Mohamed Alrifai for their constant support.

I am also grateful to Mr. Jonathan Tomalty and Ms. Kristin Berg who patiently edited my thesis writing.

I also appreciate the financial support of the Natural Science and Engineering Research Council of Canada (NSERC) and EPCOR, (an Edmonton-based utilities company) who also providing the required data.

## Table of contents

Abstract.....	ii
Acknowledgments.....	iv
Table of contents.....	v
List of Tables .....	viii
List of figures.....	ix
Chapter 1: Introduction.....	1
1.1 Background and motivation .....	1
1.2 Research objectives .....	2
1.3 Thesis organization .....	3
Chapter 2: Literature review .....	4
2.1 Overview .....	4
2.2 Background .....	4
2.3 Types of Sewer system.....	7
2.4 Assessment techniques of underground Pipes .....	8
2.4.1 Smoke testing .....	9
2.4.2 Dye Testing.....	10
2.4.3 Zoom Camera Technology .....	11
2.4.4 Sewer Scanners and Evaluation Technology (SSET).....	11
2.4.5 Ground penetration Radar (GPR).....	12

2.4.6	Ultra-Sonic Inspection (Sonar).....	12
2.4.7	Laser Interferometer .....	13
2.4.8	Infrared Thermography.....	13
2.4.9	Impulse Hammer .....	14
2.4.10	Closed Circuit Television (CCTV):.....	15
2.5	Image registration.....	18
2.6	Sewer pipe deterioration models .....	24
2.6.1	Statistical Models.....	25
2.6.2	Artificial Intelligence (AI) Models .....	28
Chapter 3: Methodology .....		33
3.1	Overview .....	33
3.2	Framework for detection of defect development .....	35
3.2.1	Input data .....	36
3.2.2	Output of the framework: .....	38
3.3	Experiments of AS method .....	39
3.3.1	Area scaler (AS).....	39
3.3.2	Experiment AS concept on CCTV Images: .....	40
Chapter 4: Implementation process of the framework.....		54
4.1	preparation and preprocessing.....	56
4.1.1	Data preparation .....	56

4.1.2	Gray scale conversion.....	57
4.2	ORB Feature detection and matching: .....	59
4.2.1	FAST (Features from Accelerated Segment Test) Corner detection Algorithm: ...	60
4.2.2	BRIEF (Binary Robust Independent Elementary Features): .....	65
4.2.3	Feature Point matching .....	68
4.3	defect segmentation and change calculation .....	69
4.4	Case studies .....	70
4.4.1	Case study (1) .....	70
4.4.2	Case study (2) .....	73
Chapter 5: Conclusion and future works .....		77
5.1.	General conclusion.....	77
5.2.	Research contribution.....	78
5.3.	Research limitation.....	78
5.4.	Recommendation and future work .....	79
References.....		80

## List of Tables

Table 2. 1: Canada's municipal infrastructure stock (2000) .....	4
Table 2. 2: Age of Canada's infrastructure (years).....	5
Table 2. 3: Cost comparison of a number of commonly used inspection methods .....	6
Table 2. 4: List of CCTV Equipment inside a CCTV Truck (Navab-kashani 2014) .....	17
Table 3. 1: AS calculation for the first pair of images.....	44
Table 3. 2: Results summary of AS experiment for the first pair of images .....	45
Table 3. 3: AS calculation for the first pair of images.....	47
Table 3. 4: Results summary of AS experiment for the second pair of images.....	48
Table 3. 5: The results of applying the concept of AS on 49 different pairs of images. ....	<b>48</b>
Table 4. 1: The calculation of first value of Area Scaler in the first case study .....	70
Table 4. 2: The calculation of first value of Area Scaler in the first case study .....	70
Table 4. 3: The calculation of third value of Area Scaler in the first case study .....	71
Table 4. 4: The calculation of first value of Area Scaler in the second case study .....	73
Table 4. 5: The calculation of second value of Area Scaler in the second case study.....	73
Table 4. 6: The calculation of third value of Area Scaler in the second case study .....	74

## List of figures

Figure 2. 1: Infrastructure condition as determined by maintenance(MIRZA 2007) .....	6
Figure 2. 2: Combined Sewer System (CSS) (EPA, 2004).....	7
Figure 2. 3: Separate Sanitary Sewer and Storm Sewer System (SSS) (EPA, 2004).....	8
Figure 2. 4: Pipeline Inspection Methods (Allouche and Freure 2002).....	9
Figure 2. 5: Typical Smoke Testing Operation(Allouche and Freure 2002) .....	10
Figure 2. 6: Dye Testing after (Khazraeializadeh 2012).....	10
Figure 2. 7: SSET Survey Machine and Some Typical Output .....	11
Figure 2. 8: Locating Pipelines using Ground Penetrating Radar Technology .....	12
Figure 2. 9: Ultrasonic Inspection for Cracks.....	13
Figure 2. 10: Infrared Thermography Inspection (www.maverickinspection.com) .....	14
Figure 2. 11: Impulse Hammer and Accelerometers (www.used-line.com)(Mellett 1995) .....	15
Figure 2. 12: CCTV crawler(Navab-kashani 2014).....	15
Figure 2. 13: Examples of defects and tap in sewer pipes(Yin et al. 2020).....	16
Figure 2. 14: Working Area inside a CCTV Truck .....	18
Figure 2. 15: Registration process main steps .....	19
Figure 2. 16: Image registration applications .....	20
Figure 2. 17: Classification of Sewer Deterioration Models .....	25
Figure 3. 1: Overview of the automated defect development detection system .....	36
Figure 3. 2 Time-series sewer pipe defect image in 2008 and 2010.....	37
Figure 3. 3: Image to the left manually bounding the defect using Microsoft PowerPoint software, the image in the right YOLO defect detection and bounding .....	38
Figure 3. 4: Illustration of the AS factor.....	39

Figure 3. 5: A flowchart of AS experiment process .....	42
Figure 3. 6: a) A pair of matching images from pipe segment 30796, b) The first pair of matching areas, c) The second pair of matching areas, d) The third pair of matching areas .....	43
Figure 3. 7: A plot of the selected points in each image.....	44
Figure 3. 8: a) A pair of matching images from pipe segment 53003, b) First pair of matching areas, c) The second pair of matching areas, d) The third pair of matching areas .....	46
Figure 3. 9: A plot of the selected points in each image.....	47
Figure 3. 10: Maximum recorded variation in AS value over a pair of images .....	51
Figure 3. 11: Maximum variation in AS taking the average of 3 iterations .....	52
Figure 3. 12: Comparison between “Difference Range” and “Difference Range (Average)” .....	53
Figure 4. 1: Implementation of framework process.....	55
Figure 4. 2: Flow chart of input data preparation .....	56
Figure 4. 3: Applying different grey scaling methods, 1 coloured image, 2 average grey scaling method, 3 lightness grey scaling method, 4 luminance grey scaling method .....	58
Figure 4. 4: Flow chart of Image matching process based on the ORB algorithm (Luo et al. 2019) .....	60
Figure 4. 5: Corner detection test (Roston and Drummond 2006) .....	61
Figure 4. 6: A pyramid or multiscale representation of a single image .....	64
Figure 4. 7: Implementation of the framework in the first case study .....	72
Figure 4. 8: Implementation of the framework in the second case study .....	75

## **Chapter 1: Introduction**

### **1.1 Background and motivation**

Human societies depend greatly on drainage systems in order to maintain a high level of public health and hygiene, which constitute a fundamental in terms of avoiding large-scale epidemics (Meeker 1971). Generally, municipalities or private utility companies are responsible for installing and/or maintaining underground infrastructures such as drainage or water supply systems. Because buried pipes are subjected to mechanical stress (e.g., vibrations from traffic, tree roots, ground settlement) and aggressive chemical environments (e.g. corrosive gases, chemical aggregation of particles), they can age at a rapid pace. Around the world, the age of existing underground infrastructure can be as old as the city itself. The burden on municipalities to prioritize and maintain rapidly deteriorating sewer pipelines is increasing (Chughtai and Zayed 2008). A preventative maintenance practice aims at detecting as early as possible the structural and operational deterioration of sewer pipes in order to take the most appropriate corrective actions, e.g., repair, rehabilitation, or if the damage is extreme, replacement of the faulty sections( Yin et al. 2020). In practice, a wide range of pipeline inspection techniques is used to investigate the structural and operational status of pipes from which deterioration models are built in order to predict their rate of failure. Presently, one of the most widely used practices for conducting sewer pipe inspections is closed-circuit television (CCTV) (Daher 2015) because this technology is reliable, robust, easy to set up (in the field), and more importantly because it allows sewer technologists to review the collected data (in the form of video footage) anytime if needed. In practice, the use of CCTV videos is generally limited to identifying and classifying defects according to some standardized code that can be internal to the organization or as defined by a professional association. However, despite the widespread of CCTV technology for sewer pipe

inspection, only limited efforts have been devoted to extracting additional information that can benefit longitudinal defect modelling and maintenance planning. Such information can be detected using time-series images of the same pipe and can be employed to gain a deeper understanding of the deterioration process in underground pipes and build more robust deterioration models.

## **1.2 Research objectives**

This research introduces an image-based framework that takes advantage of the time series of CCTV inspection videos in order to extract information conveyed by more than one image. More specifically, the output of the proposed framework is the relative geometrical development of specific defects over the service period of the sewer pipes, which addresses a significant research gap in the preventative maintenance of sewer pipes. Such deterioration information is valuable in terms of gaining a deeper understanding of how sewer defects evolve locally (over time), so as to both schedule the appropriate maintenance that will prevent the collapse of a sewer pipe and to build more robust deterioration models. The objectives of this research are outlined as follows:

- a) exploring the possibility of extracting additional value by stacking images in the form of a time series allowing time-dependent data-driven deterioration modelling to be conducted;
- b) proposing a framework for calculating the relative growth of a defect in an underground pipe;
- c) integrating image processing algorithms in order to automate the process of image registration in the field of underground pipe; and
- d) using case studies to validate and implement the framework on a dataset of real sewer pipe images.

### **1.3 Thesis organization**

This thesis is organized into five chapters starting with the introduction in Chapter 1, which introduces the topic and research objective and provides an overview of the research. Chapter 2 provides a literature review of the methods used in underground condition assessment, the process, and the application of image registration and the previous studies of underground pipes deterioration models. Chapter 3 comprises an overview of the proposed research methodology, explaining the main parts of the defect development detection system, and illustrating the area scaler concept. Chapter 4 includes an explanation of the implementation of the proposed system as applied to case studies. Finally, Chapter 5 covers the general conclusion, along with research contributions and limitations.

## Chapter 2: Literature Review

### 2.1 Overview

This chapter reviews research studies of relevance to the present research. First, background information is provided in the form of statistics and an overview of infrastructure systems in Canada, including the contribution of sewer systems and their associated costs for inspection and maintenance. Next, a literature review of various inspection and condition assessment techniques is presented with a particular focus on the CCTV process, which is the input data source of the present research. Thereafter, the image registration process is explained in terms of the main steps and classification depending on the application domain, in addition to a review of previous research pertaining to image registration employment. Finally, the literature regarding deteriorations in the sewer pipes is reviewed, with a focus on the models used to evaluate and predict the condition of the underground pipes.

### 2.2 Background

In Canada, most of the infrastructure is reaching the end of its service life. The majority of this infrastructure and facilities are managed and owned by municipalities and were constructed between the 1950s and 1970s, which means they are on the verge of failure. Given the tight fiscal situation in most municipalities, the quality of services provided is compromised (Mirza 2007). Table 2.1 shows the proportion of different infrastructure in Canada that is managed by municipalities as of 2000.

**Table 2. 1: Canada's municipal infrastructure stock (2000) (Mirza 2007)**

Canada's municipal infrastructure stock (2000)	
Transportation and transit	55%

Water and wastewater	30%
Other infrastructure (recreation, public building)	15%

In a report for the Federation of Canadian Municipalities, Mirza (2007) categorized the infrastructure according to its age and found that infrastructure that is more than 80 years old represents about 28% of the total stock, while infrastructure that is between 40 and 80 years old represents 31% of the stock. The remaining 41% per cent of Canadian infrastructure is 40 years old or less. Table 2.2) shows the percentage of the infrastructure classified in each age group.

**Table 2. 2: Age of Canada’s infrastructure (years) Mirza (2007)**

Age of Canada’s infrastructure (years)	
0–40 years	41%
40–80 years	31%
80–100 years	28%

Drainage systems in Canada are not exempt from aging and rapid deterioration as approximately 80% of sewer systems have reached the expected service life of 50 years (Zhao 1998). Mirza (2007) has found that the deterioration rate is generally correlated with the maintenance level. Proper and adequate maintenance increases the service life of the infrastructure. Figure 2.1 shows the effect of four different levels of annual maintenance on the infrastructure condition. It is widely understood that in the absence of maintenance, the deterioration rate will increase rapidly, and the infrastructure/facility will be out of service in a relatively short time. However, the deterioration rate can be lowered, and a longer service life can be achieved by investing 2% of the annual facility cost in its maintenance.

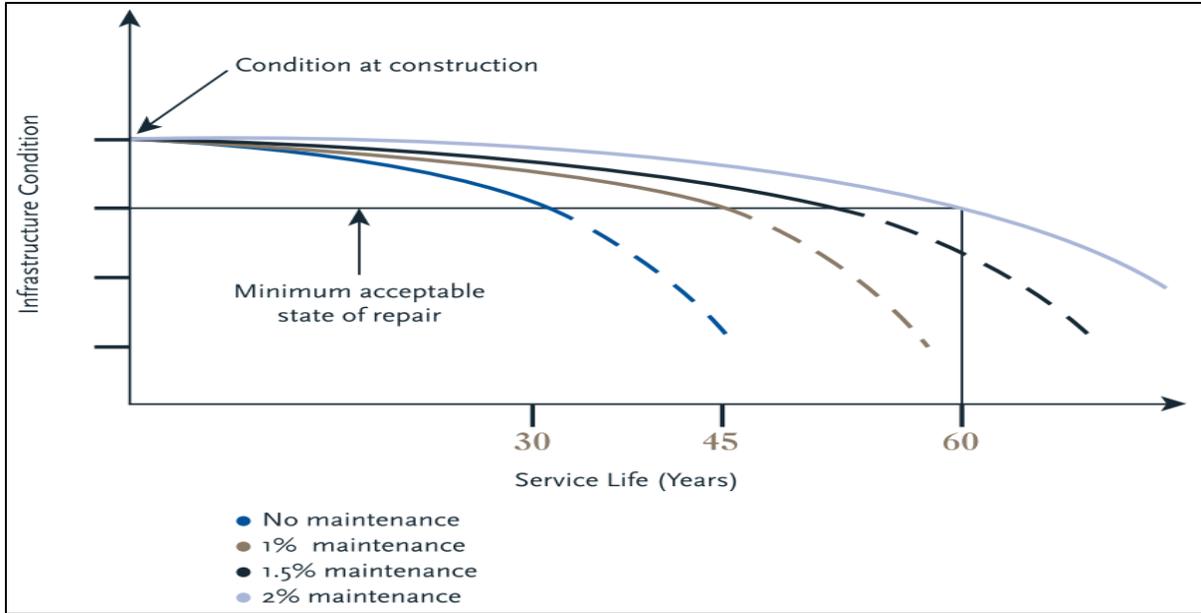


Figure 2. 1: Infrastructure condition as determined by maintenance(Mirza 2007)

Pipe inspection is one of the essential steps in assessing the condition of the pipes in the network and determining which parts need to be maintained before a catastrophic failure occurs. Table 2.3 compares the costs for some commonly used inspection methods. Inspection methods have been evolved significantly over the past 50 years. Data collected by pipe inspection has improved in terms of quality and quantity due to the updating of methods used and by incorporating the latest technologies.

**Table 2. 3: Cost comparison of several commonly used inspection methods for underground pipes(Zhao 1998)**

Method Used	Unit Cost (\$/mm diameter)
CCTV	\$0.009
Sonar	\$0.030
Man-Entry	\$0.007
Combined CCTV/Sonar	\$0.013
Zoom Camera	\$0.033

## 2.3 Types of sewer systems

There are two main types of urban drainage systems: the combined sewer system (CSS) and the separate sewer system (SSS) (Toffol et al. 2007). In CSS, waste water from domestic, commercial, and industrial uses and storm water are collected and run in a single pipe to a publicly owned treatments works (POTW). Mohammadi (2019) mentioned that CSS is convenient when the city needs both sanitary and storm sewers as it is less expensive compared to constructing two separate systems. Large cities tend to have combined drainage systems in order to help with flood control.

Figure 2.2 shows an illustration of a combined sewer system.

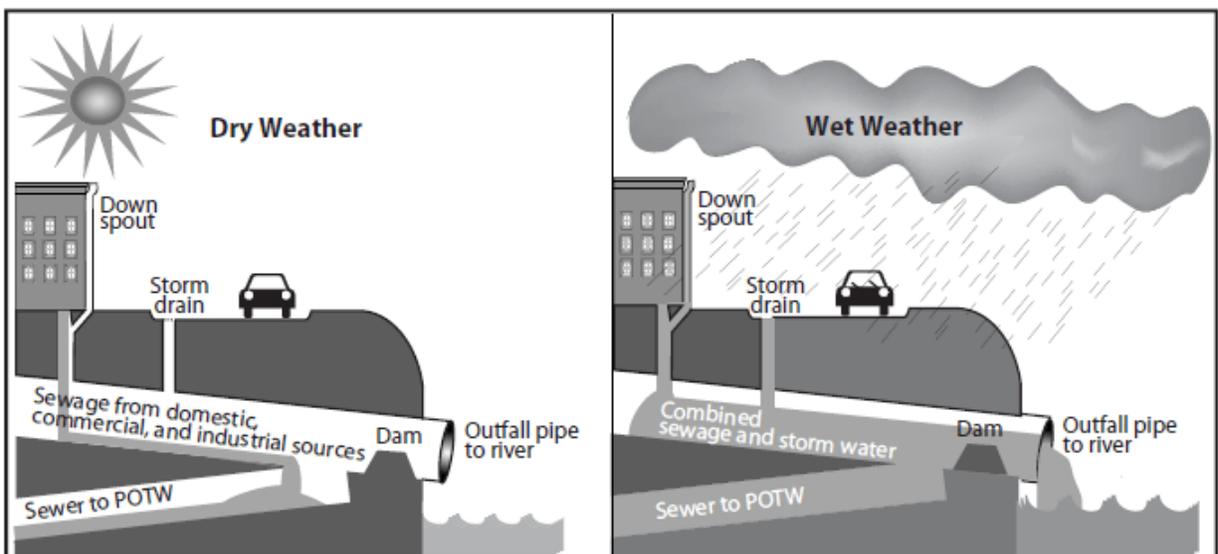


Figure 2. 2: Combined sewer system (CSS) (Mohammadi 2019)(EPA, 2004)

SSS depends on independent pipes that are completely separated from each other. One of the pipes transports sanitary water from different uses and the other transports the stormwater. This type of system is less expensive for municipalities that desire only wastewater collection; however, this system is not designed to manage a large amount of water from wet weather events as there is a separated network for dealing with the stormwater. Figure 2.3 shows a drawing of a SSS system (EPA, 2004).

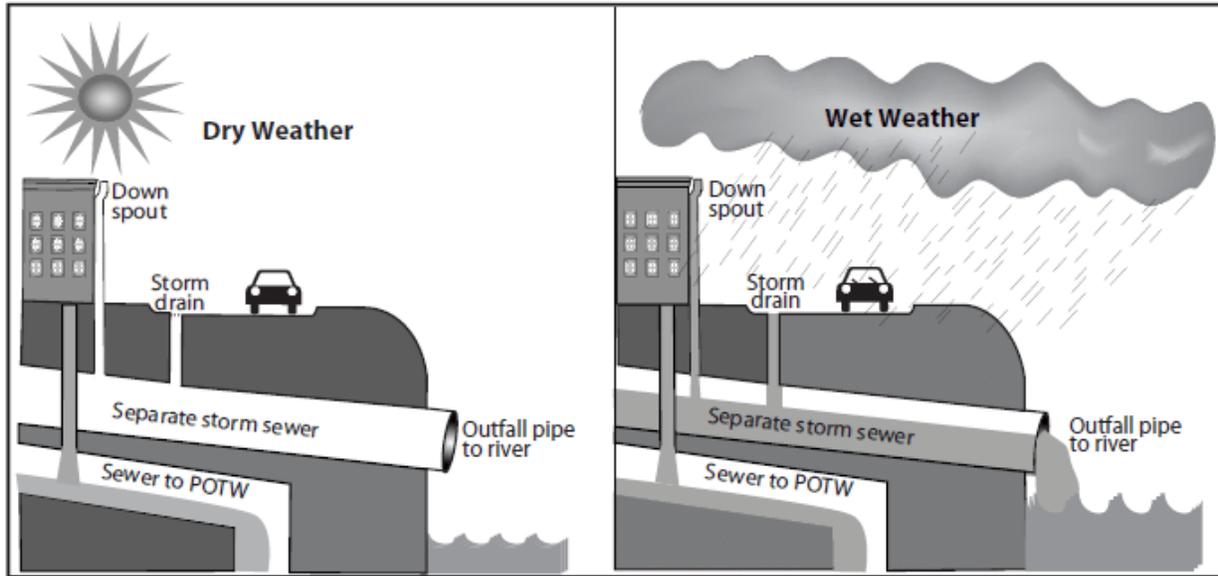


Figure 2. 3: Separate Sanitary Sewer and Storm System (SSS) (Mohammadi 2019) (EPA, 2004)

## 2.4 Assessment techniques of underground Pipes

Wastewater infrastructure systems are subject to a combination of physical and chemical factors that cause deterioration over time. A preventative maintenance practice aims at detecting as early as possible structural and operational deteriorations of sewer pipes in order to take the most appropriate corrective actions, e.g., repair, rehabilitation, or if the damage is extreme, replacement of the faulty sections (Yin et al. 2020). The pipes inspection process helps in reducing the cost of emergency repairs, which are 50% more expensive than repairs under normal circumstances (Mohammadi 2019). Interest in buried pipeline inspection first emerged after World War II (Allouche and Freure 2002). Of the various inspection methods, the man-entry method, considered the most basic inspection method, is a method in which trained personnel physically enter the pipe, performs the inspection, and records the notes. However, safety issues, the high cost, and the physical constraints, such as the diameter of the pipe that allows the inspector to enter and perform the inspection, make this technique less popular compared to other methods. Factors such as the

pipe’s diameter, age, material, and the service it delivers are important factors that determine the suitable inspection method among those shown in Figure 2.4.

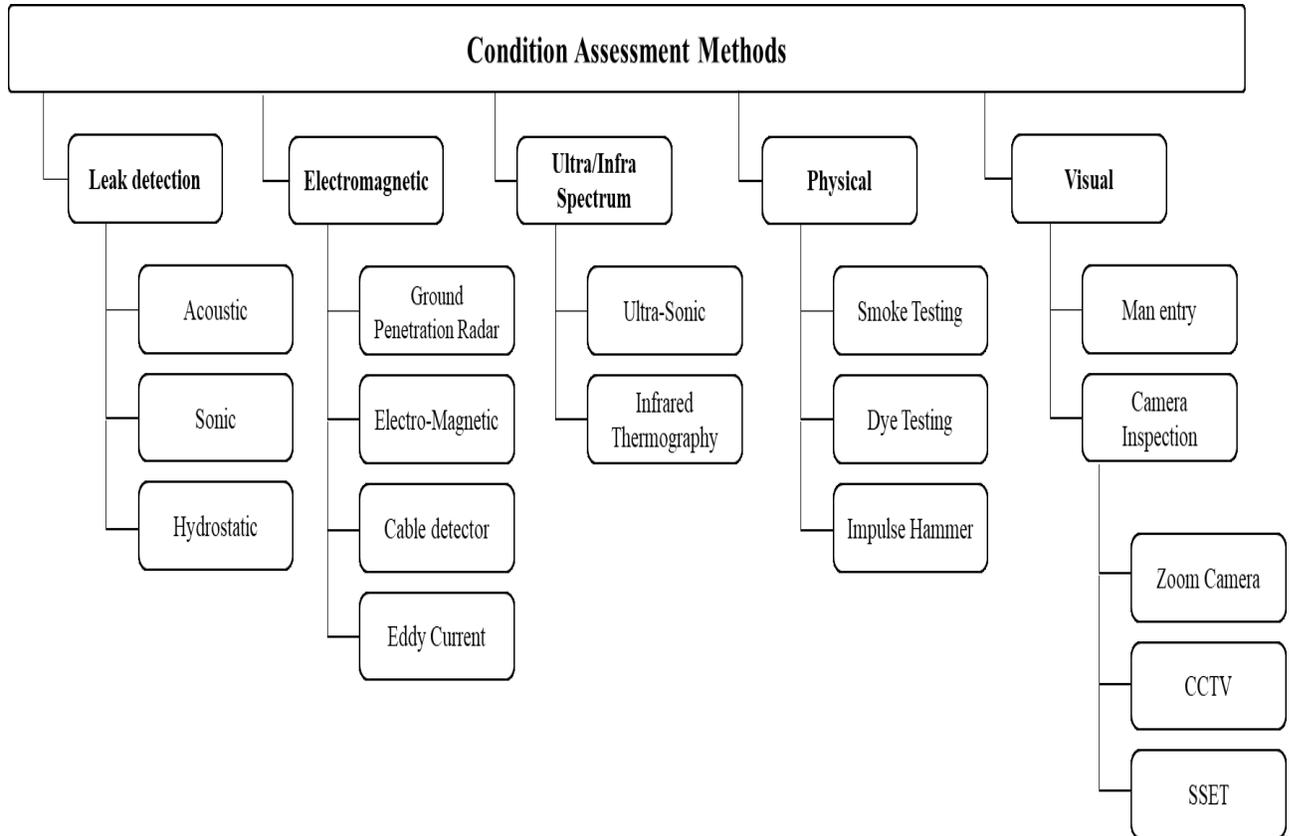


Figure 2. 4: Pipeline inspection methods (Allouche and Freure 2002)

The following is a brief description of the various inspection methods commercially used for the condition assessment of wastewater pipes:

### 2.4.1 Smoke testing

One effective tool for leak detection is smoke testing as it is a fast and low-cost technique that involves generating non-toxic smoke and pushing it into the system using a blower installed on a manhole when a part of the network has been isolated by rubber flanges or other stoppers to confine the smoke. As a result, smoke escapes from faulty connections, or cracks, which allow the detection of these defects. Figure 2.5 shows the smoke testing process.

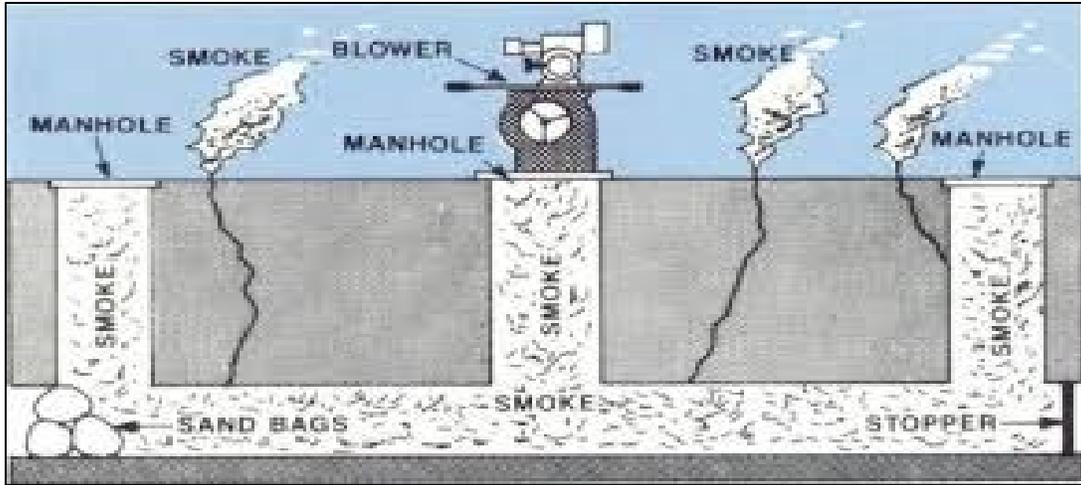


Figure 2. 5: Typical smoke testing operation (Allouche and Freure 2002)

### 2.4.2 Dye testing

In the dye testing method, a non-toxic dye is used to detect any unrecorded connections between the stormwater drains and sanitary sewers. The non-toxic dye which has a fluorescent colour is applied in the area of interest, for example, through drain leaders, driveways, or local drains. Then, a manhole located downstream from the area of dye application is monitored, as shown in Figure 2.6, to detect the trace of the dye (Allouche and Freure 2002; Khazraeializadeh 2012).



Figure 2. 6: Dye testing (Khazraeializadeh 2012)

### 2.4.3 Zoom camera technology

A visual inspection is performed by using a high-tech camera equipped with a long-range zoom lens. The interior of the pipe is illuminated by means of powerful halogen lights and the camera provides a continuous image of the pipe's interior surface. This method depends on the penetration depth of the lights which can achieve an operational range of a zoom camera with a range of 25m in pipes that have a diameter of 250 mm or less, or a maximum range of 50m in a pipe with a larger diameter (Allouche and Freure 2002).

### 2.4.4 Sewer scanners and evaluation technology (SSET)

Sewer scanners and evaluation technology (SSET) provides a continuous, 360-degree view of the interior wall of the pipe through a high-resolution scanner. One advantage of this method is producing unfolded image (2D image) that improve the accuracy of the defect's detection (Koo and Ariaratnam 2006)(see Figure 2.7). Another advantage of SSET that it can achieve better productivity as there is no need for frequent stops at the defects. (Iseley et al. 1997; Khazraeializadeh 2012;).

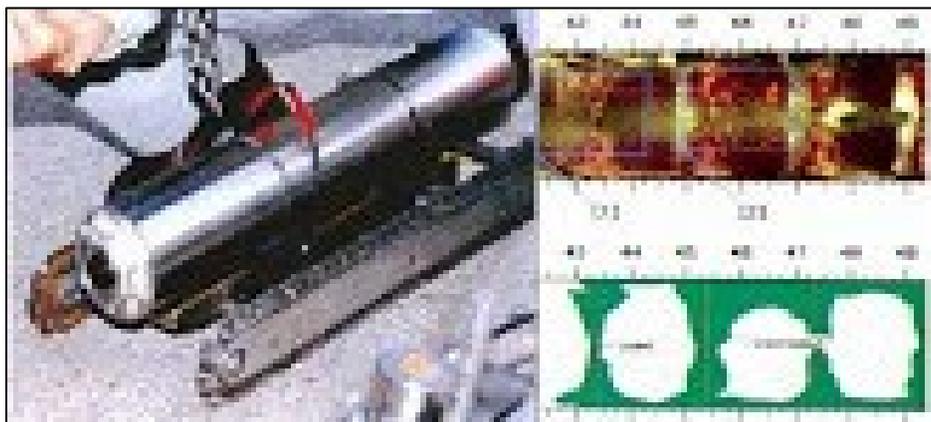


Figure 2. 7: SSET survey machine and an example of typical output (Blackhawk Inc., 2000; Allouche and Freure 2002)

## 2.4.5 Ground penetration radar (GPR)

In this method, radio waves are sent into the ground in the area of interest. Travel time and the strength of the reflected signals of the short pulses of the electromagnetic energy are received and recorded (Duran et al. 2002). GPR is an easy to set up, effective method in locating the underground pipe and identify any gaps or voids surrounding the pipe. Recently GPR has been used to evaluate the rehabilitation by installing the GPR in a crawler and sending it through the pipe. Conductive material such as clay minerals can negatively affect the accuracy and the depth of penetration of this method ( Khazraeializadeh 2012; Mellett 1995).

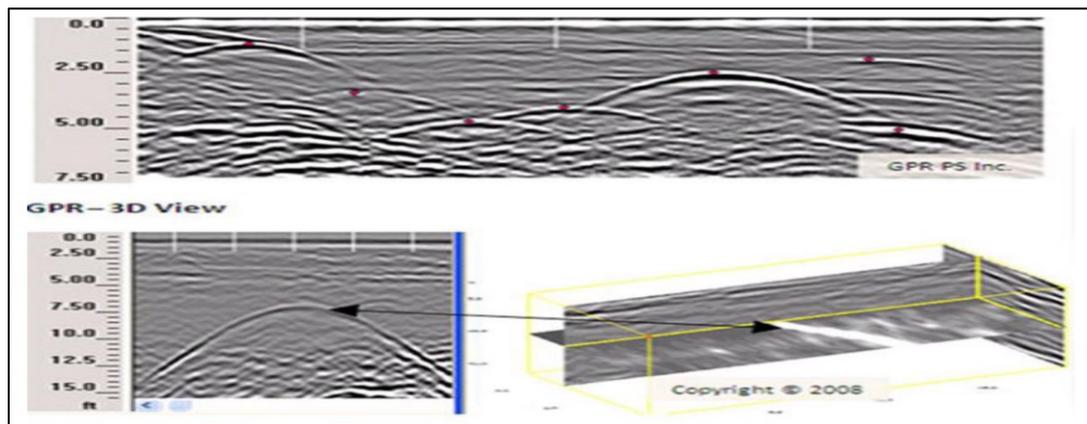


Figure 2. 8: Locating pipelines using ground penetrating radar technology  
(www.gprps.com)(Khazraeializadeh 2012)

## 2.4.6 Ultra-sonic inspection (sonar)

Ultra-sonic or sonar inspection is a non-destructive method that use high frequency sound waves to inspect the interior wall of a pipeline(Khazraeializadeh 2012). The waves are directed to the pipe and the differences in the amount of the energy reflected and the travel times of these waves are used to calculate the depth of the object and identify its location. Ultra-sonic inspection has higher Capability of detecting fine Defects such as small size cracks and fractures in addition to pipe-wall deflection and displacement(Koo and Ariaratnam 2006)(see Figure 2.9). In addition,

information about the structure of the pipe, such as thickness, and the material in the surrounding area of the pipe can be obtained (Allouche and Freure 2002).

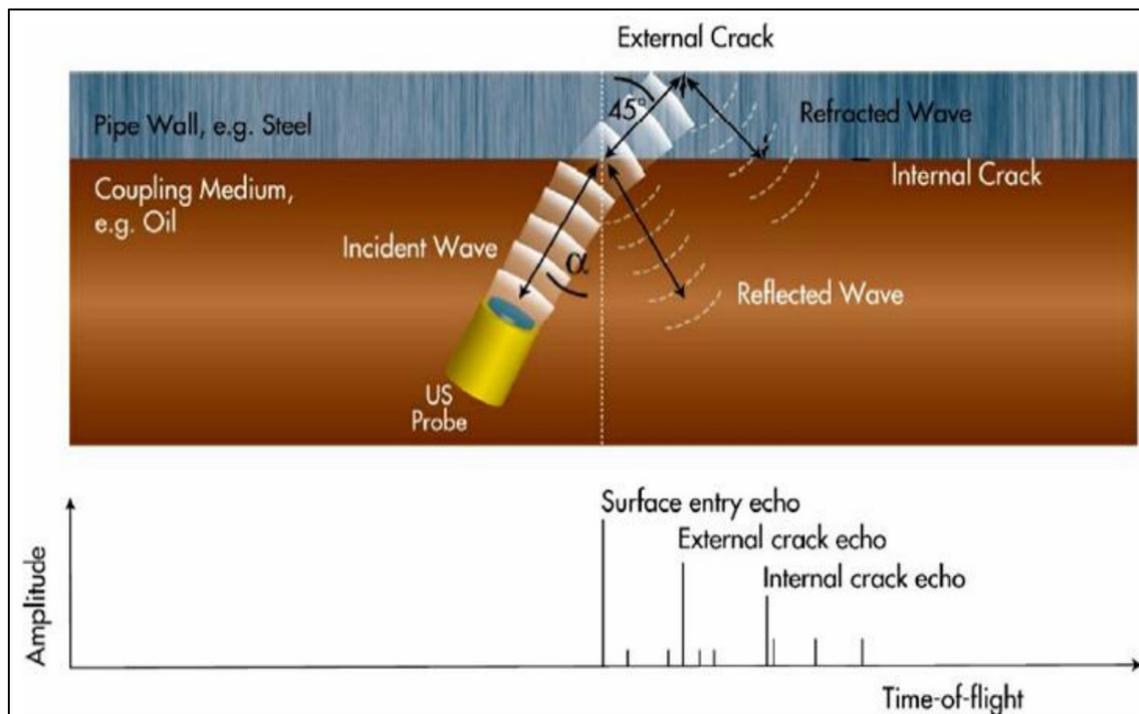


Figure 2. 9: Ultrasonic inspection for cracks  
(www.ppsa-online.com)(Khazraeializadeh 2012)

#### 2.4.7 Laser interferometer

The laser interferometer method works by way of a similar mechanism to the sonar method. However, in a laser interferometer, a laser light will be directed toward the pipe, a surface with defects will reduce the amount of light reflected. The 3D shape of the pipe and the defects are obtained, which helps in locating any defects so accurate measurements can be obtained (Allouche and Freure 2002).

#### 2.4.8 Infrared thermography

Infrared thermography relies on the principle that different objects have different thermal conductivities; thus, the flow of heat energy will vary in terms of speed when it moves between

objects with different temperatures (from warm areas to cooler ones). An image is obtained that shows different colours representing the different objects depending on their level of infrared radiation (see Figure 2.10). Infrared thermography is a safe method in terms of using only thermal radiation (Duran et al. 2002). Leaking points and voids around the pipe can be effectively detected; however, the exact location and the extent of the defects cannot be defined accurately (Khazraeializadeh 2012).

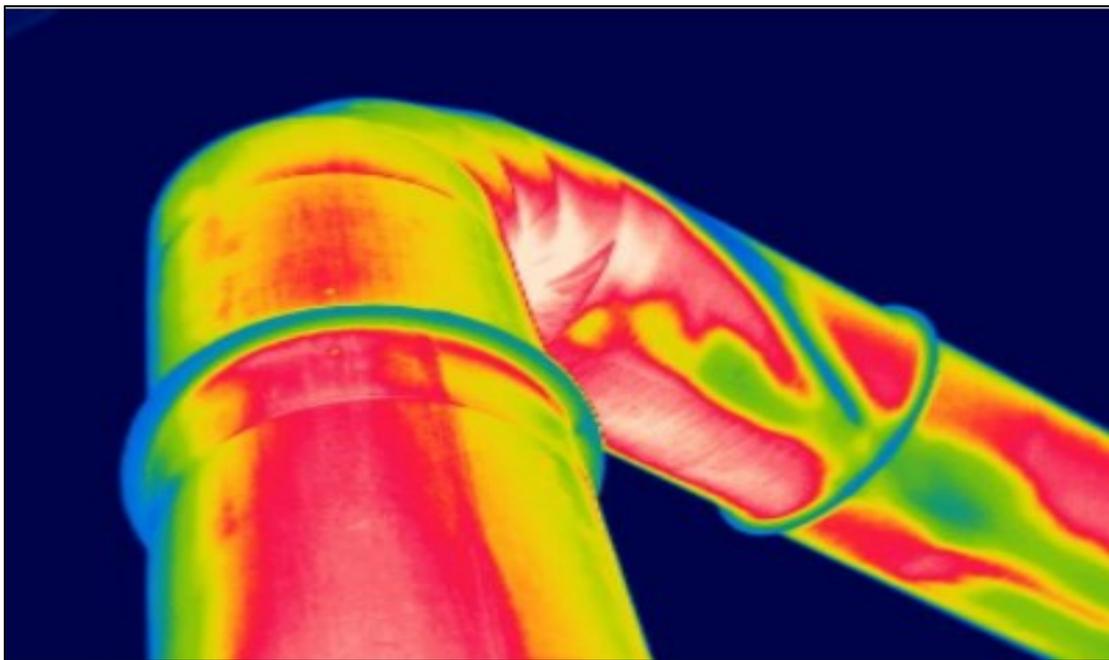


Figure 2. 10: Infrared thermography inspection(Khazraeializadeh 2012)  
([www.maverickinspection.com](http://www.maverickinspection.com))

#### **2.4.9 Impulse hammer**

The impulse hammer method involves monitoring the response of the inspected sewer to the dynamic broadband frequency generated by a dynamic hammer placed at a manhole (see Figure 2.11). The impulse hammer method is mainly used to inspect the structural soundness of brick sewers (Khazraeializadeh 2012; Sibbald et. al. 1995).



Figure 2. 11: Impulse hammer and accelerometers ([www.used-line.com](http://www.used-line.com)) (Mellett 1995)

#### 2.4.10 Closed-circuit television (CCTV)

The most common technique used for the inspection of sewer pipelines is closed-circuit television (CCTV)(Koo and Ariaratnam 2006). The emergence of CCTV occurred after World War II as most of the sewer systems were damaged and needed to be inspected (Allouche and Freure 2002). Although the basic concept behind CCTV is to drive a crawler equipped with a camera through the drainage pipe and the camera televises the interior image of the pipe, many improvements have been introduced to the method. The camera that is driven inside the pipe (see Figure 2.12) can be remotely directed and a zoom lens can be added to focus on detected defects.

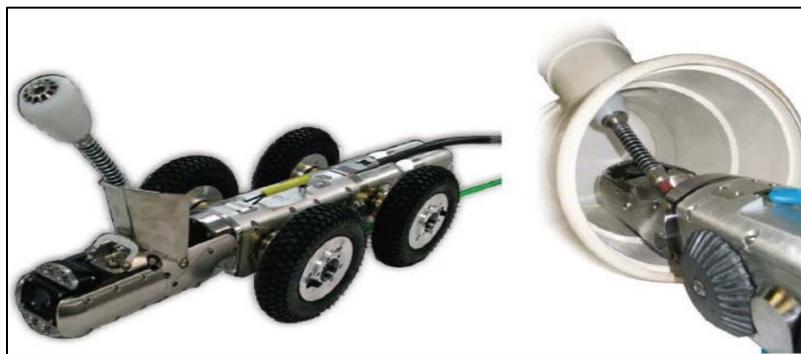


Figure 2. 12: CCTV crawler(Kashani 2014)

The visual inspection enables the observing of the defects in the pipe, however, the interpretation of the images depends on the operator, which can result in differences in the description of the detected defects provided by the different operators. There are several protocols for condition

assessment of the sewer systems that are used by operators to recognize and define defects (Kashani 2014). A wide range of defects in sewer pipes can be detected by visual inspection, such as crack, fracture, holes, and deposits (see Figure 2.13).



Figure 2. 13: Examples of defects and example of a tap in sewer pipes (Yin et al. 2020)

The CCTV process team usually consists of four members: a) one flusher operator who flushes the pipes before the CCTV inspection process, because flushing dirty pipes helps in obtaining clearer images and opens the path for the crawler; b) One CCTV personnel who performs the inspection process by driving the crawler, recording the videos and making the daily reports; c) two general swamper who help with opening the manhole, installing the crawler, organizing the traffic, and supporting the CCTV process in general. A CCTV unit is a modified truck to support the process that is equipped with all the needed tools. Table 2.4 shows the various tools typically contained in a CCTV truck. The truck also has space to do the office work and is equipped with a computer and a printer to prepare the daily reports (see Figure 2.14).

**Table 2. 4: List of CCTV equipment inside a CCTV Truck (Navab-kashani 2014)**

No.	Tools/Equipment Description	Quantity
1	Canvas trap	1
2	Traffic cones	8
3	Safety signs (Men working)	2
4	Snap ring pliers	1
5	Yellow poly rope	1
6	25 pc hex key set	1
7	5-piece pliers set	1
8	7" strait jaw visa grip	1
9	Cutter wire stripper	1
10	Utility knife & blades	1
11	6", 10" and 12" adjustable wrench	1
12	12-piece screwdriver set	1
13	Nut driver 5/16"	1
14	Tape measure	1
15	Pick	1
16	Sledgehammer	1
17	2-way mallet	1
18	Multimeter	1
19	Heat gun	1
20	Walkie-talkie radios	2
21	Gas detector	1
22	Measuring heel	1



Figure 2. 14: Working area inside a CCTV truck (Navab-kashani 2014)

## 2.5 Image registration

Comparing images of the same scene captured at different times from different viewpoints can be a challenging task for algorithms not using any manual guidance. The images need to be aligned at the pixel level (i.e. registered) so that differences can be detected (Kuppala et al 2020). A means of correspondence should be established between images of interest in order to have a meaningful comparison (i.e. difference). There is no one method that can fit all cases of image matching. That is because of the diversity of degradation and distortion in images to be compared. Every method should consider not only the assumed type of geometric deformation between the images but also noise corruption and required accuracy (Zitová and Flusser 2003). In this respect, when it comes to the registration of two or more images for the purpose of pixel-based comparisons and differencing, Song et al. (2014) and Zitová and Flusser (2003) divided image registration onto four main steps as shown in Figure 2.15:

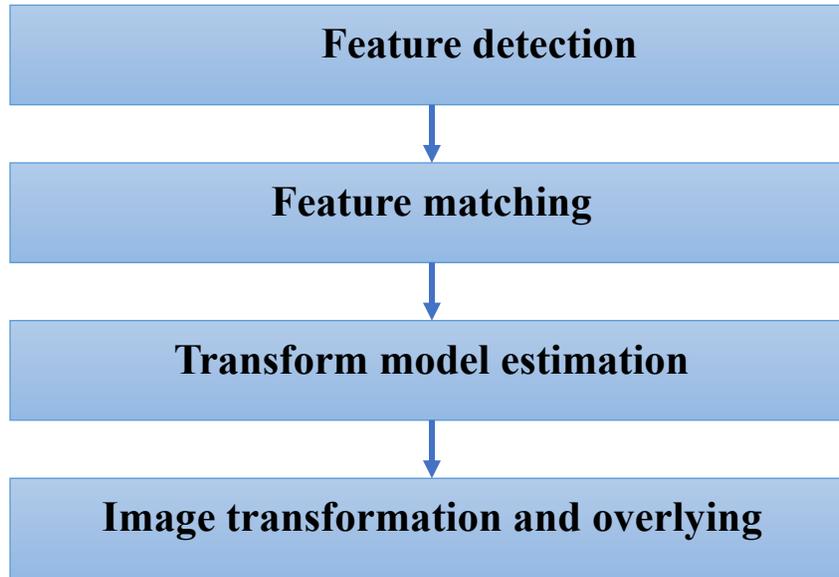


Figure 2. 15: Registration process main steps (Wyawahare et al. 2009)

- Feature detection: Distinct points that represent the salient elements in the image are detected (Siddharth and Rajeev 2014). These points can usually be used as control points. Examples of features could be lines intersections, corners, contours.
- Feature matching: When features are detected, they are usually described based on the designs and patterns of their surrounding pixels. Different algorithms use different formulas to assign a signature of a feature point (Siddharth and Rajeev 2014). One example of these descriptors is calculating a bit string description of an image patch constructed from a set of binary intensity tests for a set of sampled pairs of pixels that surround the feature point. This process is significant as it is considered the foundation of features matching. Various similarity measures then can be applied to the values of the descriptors in order to find the matching points.
- Transform model estimation: in this step, the parameters that would account for the various deformations in the image are estimated (Deshmukh and Bhosle 2011). These parameters are called transformation model parameters. Transformations are classified into three

categories: 1) Rigid transformations which include scaling, Rotation, and translation. In a rigid transformation, the overall relationships between points do not change. 2) Affine Transformations that are restricted to preserve collinearity and allow besides rotation and translation also scaling and shearing 3) Non-rigid transformations in which objects in the images do not retain their relative shape and size and the overall relationships between points change. perspective transformation is an example of a non-rigid transformation. (Brown 1992)

- Image resampling and transformation: In this step, the two images are overlaid on each other by transforming one of them depending on the estimated transformation model (Song et al. 2014).

A classification of four categories of image registration is introduced by Zitová and Flusser (2003) depending on the application, as shown in Figure 2.16.

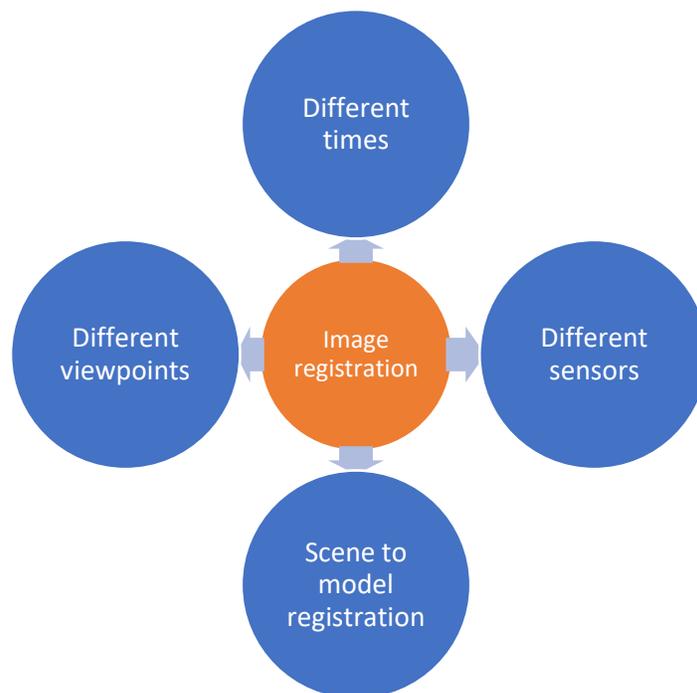


Figure 2. 16: Image registration applications

- Different times: This application is used when changes during the time of an object or movement of the object is required to be detected. Thus, the information exists in more than one image taken at different times. Time-series images acquired for the same view are required(Zitová and Flusser 2003). Some examples of this application are detecting the growth of tumors in the medical field(Oliveira and Tavares 2014), motion detection in computer vision, and changes in land usage in remote sensing fields.
- Different viewpoints: Changing the sensor's (e.g. camera) viewpoint is particularly important when several images need to be assembled to mosaic a large area or to build a 3D representation of an object.
- Different sensors (multimodal analysis): This is used when different sensors are used to get images for the same view in order to build a complex model for the view depending on the information flows acquired by the different sensors. This is widely used in aerial or satellite remote sensing.
- Scene to model registration: the aim of this type of registration is to register images of the view with a model built for the same view in order to compare them. A model, for example, can be GIS layers and the image is an aerial image taken for the same area(Zitová and Flusser 2003). This type is widely used for remote sensing to localize images, such as aerial photos, on the map, and making comparisons.

Image registration has been applied in different research areas. Song et al. (2014) proposed an image registration approach for remote sensing in which the authors employ a histogram of triangle area representation (TAR) and feature matching strategy. In their research, the approach is based on a robust transformation parameter estimation algorithm called the histogram of TAR sample consensus, which was developed to replace the existing random sample consensus (RANSAC) and

progressive sample consensus (PROSAC) methods that have been widely used in the transformation parameter estimation step of remote-sensing image registration. Zelinski et al. (2018) presented a non-rigid image registration algorithm for remote sensing that uses publicly available tools such as python, NumPy, SciPy, OpenCV, and SIFT to find anomalies associated with human activity for the purpose of detecting underground nuclear explosions. A good deal of research in the area of image registration has been done in the medical field (Oliveira and Tavares 2014). For example, Foskey et al. (2005) proposed and validated a framework, based on deformable image registration, for the automatic processing of serial 3D CT images used in image-guided radiation therapy. In the field of civil engineering, few attempts have been made to apply image registration techniques when compared with the medical and remote sensing fields. One example of those attempts is the work by Sohn et al. (2005) in which a framework using a multitemporal image of a crack was proposed to quantify the development of cracks in a concrete specimen. The images were obtained from an off-the-shelf digital camera. Two-dimensional projection based on a modified iterated Hough transform algorithm was applied to retrieve the geometric transformation parameters. The output of the system is the crack of concrete surface changes (length and width). Chen et al. (2006) studied the relationship between the expansion of concrete and crack width, in which a time series of photos was taken to examine the accelerated process of cracking generated by putting an alkali-containing concrete plate in a hot environment to produce cracks quickly. Then cracks were segmented from the images by applying a greyscale thresholding based on the observation that cracks have lower grey value than their surroundings. Another related study by Adhikari et al. (2014) proposed a framework to extract the properties of cracks in bridges using image process techniques. A Fourier Transform of digital images was used to estimate the transformation parameters in the registration process for change detection. Kim et

al. (2015) developed a crack detection system using unmanned aerial vehicles and digital image processing. The inherent structure of the crack in captured images was detected using morphology techniques. The system uses segmentation, feature extraction, and decision-making components. Yang et al. (2018) focused on detecting changes in thin cracks on reinforced concrete surfaces. In their study, several processes were integrated, namely, image-matching based on optical flow and subpixels, camera calibration to make up for the deformation caused by a change in perspective, and other required geometric transformations. The outputs of the system are detection of early-stage fine cracks that are unobservable by the naked-eye and an estimation of the width of the crack opening. An image-based framework for multiview change detection in concrete was proposed by Buatik et al. (2019) to compare images produced by 3D reference models built from a set of images of the structure with new acquired images in order to detect any changes over time.

The complexity of transformation (the third step shown in Figure 15) that is required for an image registration application for detecting changes in defects in underground pipes is a challenging task, as there are different conditions affecting the choice of the suitable transformation model. In most cases, different perspectives of the video camera caused a change in relative shape and size of objects in the images taken. Perspective distortion has two effects: 1) scaling, which causes the imagery to appear smaller the further it is from the camera; and 2) foreshortening, which causes the imagery to appear more compressed the more it is inclined away from the camera (Brown 1992).

Information regarding camera specifications and geometric measurements, such as distances from the objects, are required to eliminate perspective distortion effects. This information is not available most of the time in the context of sewer pipe inspections, especially for the distances between camera and objects, which makes the estimation of the transformation parameters a

complicated process. In fact, it is crucial to have a correspondence that is within an acceptable accuracy limit and can simplify the complexity of estimating the transformation parameters, which is difficult in underground pipes images as previously explained. When this correspondence between underground pipes is established, the information conveyed in more than one image can be retrieved, and in our case, this includes the development of defects (e.g., cracks) based on images for the same defect taken at different times.

## **2.6 Sewer pipe deterioration models**

In the case of a failure of a sewer system, there will be catastrophic consequences affecting most sectors of society. For that, wastewater is rated with zero tolerance for failure (Khazraeializadeh 2012).

To predict the health condition of sewer pipes, various deterioration models have been developed over the years. In order to understand the short-term and long-term behavior of sewer pipes, information collected from underground pipe inspections is used to build and structure the mathematical tools that can help to estimate the future condition of drainage pipes. Mohammadi (2019) classified the commonly used sewer deterioration models into two groups, statistical and artificial intelligence models, as shown in Figure 2.17.

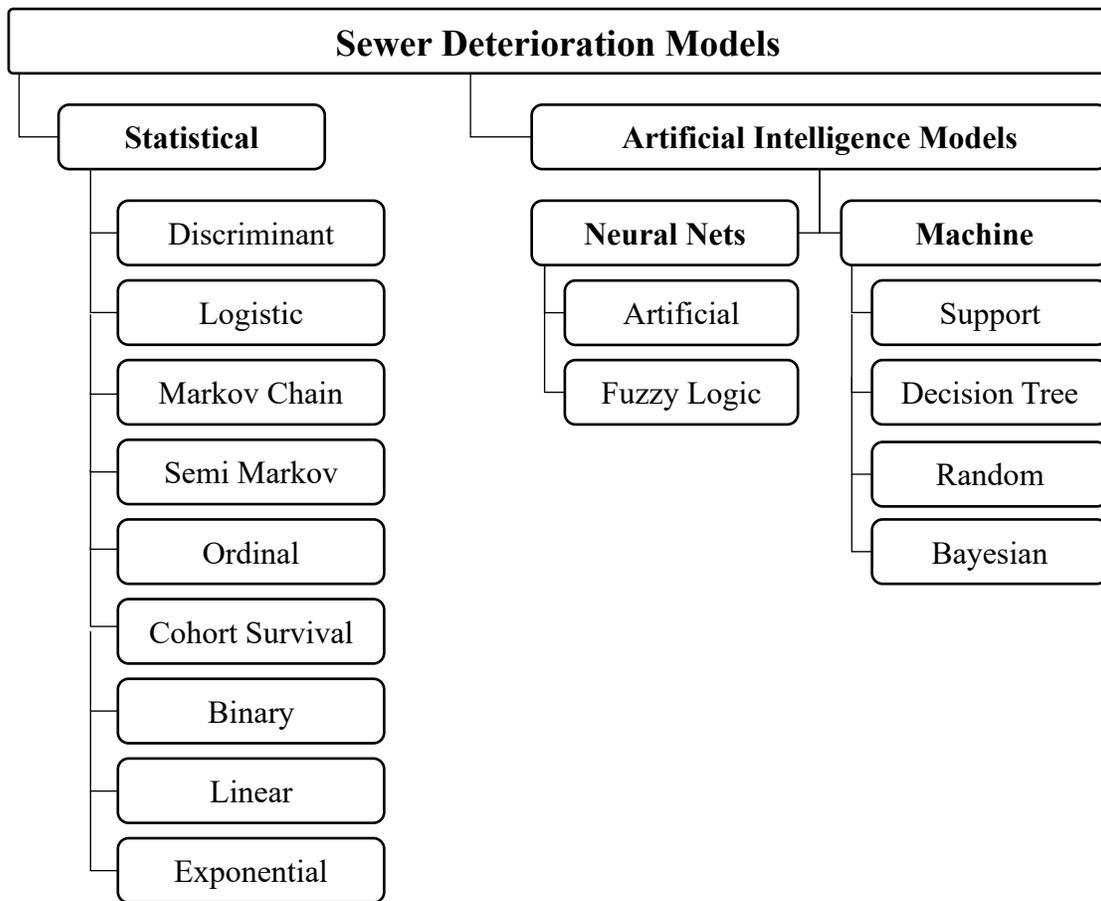


Figure 2. 17: Classification of sewer deterioration models(Mohammadi 2019)

The following is a brief description of the various types of statistical and artificial intelligence models.

### 2.6.1 Statistical models

In statistical models, a random variable  $X$  which represents a quantity, whose outcome is uncertain, is described depending on analyzing the probabilistic nature of historical data. The probability estimates which output of the statistical model can be used to predict the future condition of sewer mains (Coles 2001). The quality of data is an important factor that affects the performance of statistical models that are sensitive to noisy data according to Tran (2007). The probabilistic nature of the deterioration process makes statistical models a good choice for use in modelling sewer pipe

deterioration. A linear regression model is the simplest statistical model and it depends on predicting the value of a dependent variable by mean of an independent variable. Multilinear regression models are more complicated in that several independent variables are combined in a linear model to estimate the value of a dependent variable(Li et al. 2019). Exponential regression models, which are nonlinear models, are more flexible than linear models and can be better fitted with fewer parameters to represent the response between the dependent and the independent variables (Chatterjee and Simonoff 2013). Gedam (2016) used data obtained via closed circuit television (CCTV) inspection of a sewer system to build a linear regression model. They conclude that pipe age is the most significant factor in determining the condition of the pipe. Other factors, such as pipe installation depth, also have an impact on the regression analysis. The pipe material and pipe diameter are found to be less important. Logistic regressions are used to analyze the relationship between multiple independent variables and a categorical dependent variable (Hawari et al. 2020). El-Assaly et al. (2006) proposed a framework to estimate the cost of rehabilitation of sewer pipes in the city of Edmonton, using logistic regression deterioration to predict the pipe condition and the cost of the repair. Five factors was considered in the model; age, waste type, material type, diameter, and depth. Park (2009) categorized logistic regression models into three groups: logistic regression, multinomial logistic regression, and ordinal logistic. Rigid sewer pipe deterioration and failure was the focus of a study presented by (Davies et al. 2001). They developed a logistic regression model to predict the structural condition of rigid sewer pipes by considering the factors that affect the process of deterioration which in turn are grouped into three categories, namely, construction features, local external factors, and other factors. The result indicated that pipe material, diameter, length, sewer type, location, groundwater, and soil corrosivity are the

influencing factors that affect the deterioration of sewer pipes. Koo and Ariaratnam (2006) presented a logistic regression model to measure the extent of deterioration of the infrastructure system. They used data that was collected by ground penetrating radar (GPR) and digital scanning and evaluation technology (DSET). To generate the binary regression model pipe age, maximum velocity, and cumulative flow were used. They concluded that pipe age and cumulative flow are important factors in the process, while maximum velocity is not a main factor. A Bayesian logistic framework is developed by Kabir et al. (2018) to predict the structural condition of sewers. To validate the effectiveness of the proposed framework a total of 12,728 sewer mains belonging to the wastewater network of the city of Calgary, Canada, were used. The factors used to build the model are pipe age, material, diameter, length, slope, depth, rim elevation, up invert, and down invert. The results of the study show pipe age and length are the main factors in the deterioration of cementitious and clay sewers, whereas sewer age and diameter are enough to capture the effect for metal (cast iron, or ductile iron) and plastic sewers.

Markov chains were developed by Andrei Markov in 1906 as a discrete-time stochastic process. Markov chains provide information about the probabilities of condition changes in a system (Hawari et al. 2020). Wirahadikusumah and Abraham (2002) developed Markov chain-based models that are optimized for extending infrastructure management modelling to sewer systems. Pipe material, depth, groundwater level, and type of soil were used to generate the model. The result showed that municipalities and sanitary departments must routinely inspect their sewer systems and collect system-wide condition data in order to obtain significant improvements of deterioration models. Hyung et al. (2005) developed a Markov chain-based deterioration model to estimate the transition probabilities for a sewer system using the ordered probit model. They concluded that factors such as depth of the pipes, soil condition groundwater level, and the

frequency of sewage overflows should be collected and evaluated in order to predict more accurate and detailed deterioration patterns. Scheidegger et al. (2011) presented a novel network condition simulator (NetCoS) that produces a synthetic population of sewer sections with a given condition-class distribution. A semi-Markov chain to calculate transition probabilities based on user-defined survival functions was used in the deterioration model.

### **2.6.2 Artificial intelligence (AI) models**

The interest in employing neural networks and machine learning methodologies to build deterioration models has increased. That is because AI models have the capability to solve complex problems such as estimation of the ordinal ratings or nonlinear deterioration behaviour (Mohammadi 2019). Charniak and McDermott (1985) define AI as “the art of creating machines that perform functions that require intelligence when performed by people”. Early AI work was developed in 1943 by McCulloch and Walter Pitts based on an understanding of three components, namely, the basic physiology and function of neurons in the brain, propositional logic, and Turing’s theory of computation (Russell and Norvig 2010). In the following section, neural nets, genetic algorithms, and machine learning will be discussed briefly. A) Neural nets and genetic algorithms: The principle of neural nets and genetic algorithms is to instruct connections between entities or nodes that work like the human brain. Neural net models are used for different functions and applications. In the field of underground pipes, fuzzy set theory and neural networks (NNs) were employed to model deterioration by Tran et al. (2007), and Mohammadi (2019). Najafi and Kulandaivel (2005) developed an artificial neural network (ANN) model for predicting the condition of sewer pipes based on the historical condition assessment data. Identifying the pipes in bad condition was the output of the model and the input is a set of known values. The model showed a good learning tendency; however, the noisy data was a big problem toward gaining high

accuracy. They concluded that using ANN models for underground pipe condition prediction is feasible; however, acquiring higher accuracy requires a larger and more inclusive data set. Tran et al. (2006) presented a probabilistic neural network (PNN) to model the structural deterioration of concrete pipes. Approximately 650 data points were used to build the model, and the data were provided by the City of Greater Dandenong in Victoria, Australia. Pipe diameter, age, depth, slope, location, number of trees, hydraulic condition, soil type, and soil moisture were the input variables considered in this model. The results show that the PNN model has a better predictive power and uses significantly more input variables (i.e., explanatory factors) than the discriminant model. Khan et al. (2010) presented a study that uses artificial neural networks to investigate the importance and influence of certain characteristics of sewer pipes with respect to their structural performance. The data was provided by the municipality of Pierrefonds, Canada. Pipe material, diameter, depth, bedding material, length, and age were used to build the model. Back propagation and probabilistic neural network (NN) models were developed and validated. The developed NN models were used to rank the parameters, in order of their importance/influence on pipe condition. The developed models indicated that neural network is capable to prioritize inspection and rehabilitation plans for existing sewer mains. Sousa et al. (2014) evaluate the performance of artificial intelligence tools, namely artificial neural networks (ANNs) and support vector machines (SVMs), in predicting the structural condition of sewers. The results showed that ANNs have the potential to achieve the best results when compared to the SVMs and are sensitive to the approach used in the development and their performance may vary in a wider range of results when compared to the SVMs. However, the SVMs are also less sensitive to the independent variables used in the model. Due to overlapping in the solution space for both models, it was not possible to select the best model in this study. The superiority of one model over the other depends on the

goal of the inspection and the development method of the tool. A trade-off between lower performance and less variability (support vector machine) and higher performance and more variability (artificial intelligence) should be considered. Jiang et al. (2016) proposed ANN model to predict the corrosion processes of concrete sewer pipes. the expected start time of corrosion and the rate of development can be predicted by the developed model. Actual field measurement of Australian sewer pipelines were used to develop and validate the model. Alsaqqar et al. (2017) compared the performance of multiple discriminant model (MDM) and neural network deterioration models (NNM) to predict the condition of sewer system in the city of Baghdad. They found that NNM has higher overall prediction with an efficiency of 87.3%. Age, traffic, diameter, material, type, slope, length, and depth are considered in developing the models. B) Machine learning: Machine learning is a branch of artificial intelligence that intersects with other fields, especially statistics, mathematics, physics, theoretical computer science, and more. Gobeyn et al. (2019) reformulate the definition of machine learning as a field of study that uses computer programs or algorithms that can adapt or change (i.e. train) from an experience (data) given a performance measure. Machine learning algorithms are categorized into three broad categories based on the method of learning: according to Sathya and Abraham (2013) and Mohammadi (2019)

- Supervised learning: Learning from an experience by providing training data that includes examples of input variables and their corresponding output. The machine learns from the data provided and infers a function to get the output from the corresponding input.
- Unsupervised learning: Learning from an experience by providing training data that includes only input data without a corresponding output.

- Reinforcement learning: Like unsupervised learning in terms of providing input data without their corresponding output. It learns from the interaction with the different factors through the trial and error approach, for example, by mean of reward and penalties method.

Various machine learning algorithms, such as support vector machine (SVM), decision trees random forest, and Bayesian regressions, have been used to model the deterioration process of underground pipes. Mashford et al. (2011) investigated the use of support vector machine (SVM) models to predict the condition of sewers. CCTV footage collected of the sewer system in Adelaide, Australia, was used to build the model. The data has discrete condition grades for the inspected pipes that vary between 1 (good) and 5 (very poor). The input variables of the model were pipe diameter, age, road type, slope, start invert, end invert, material, soil type, soil corrosivity, grade, angle, sulphate soil, and groundwater level. The SVM model was tested by using this sample data before being applied to predict the condition of the remaining assets in the sewer system. The results of the study proved that the SVM achieves good predictive performance with 91% accuracy and can be used as a new approach to model deterioration of sewer pipes. Syachrani et al. (2013) developed a decision tree-based deterioration model for drainage system pipes. The study used different factors as an input of the model, namely, pipe age, diameter, length, slope, and the number of pipe defects. The data were provided by Johnson County Wastewater (JCW) in Kansas, United States. Conventional regression- and neural network-based models that use the same data sets were developed and a comparison of the performances was undertaken. The outcome reflected that the decision tree model consistently outperforms regression and neural network models. Laakso et al. (2018) developed random forest and binary logistic regression models to predict condition rating of sewer pipelines. The databases used in this research were

collected from southern Finland. The results showed the random forest model (accuracy 62%) slightly outperformed the binary logistic regression model (accuracy 56%).

As it can be seen from the literature review study, almost all the deterioration models focus on predicting the overall health condition of sewer pipes rather than the local evolution over-time of an anomaly. Previous research focused on quantifying the collective impact of defects (in the form of deterioration) and purposely overlooked the importance of local changes of defects. In fact, according to the well-known “process of sewer collapse” theory, the collapse of a sewer pipe proceeds in three stages. In the initial deterioration stage, the pipe suffers one or several minor defects evolve in size (which during the second phase is referred to as deterioration) to reach the final stage of deterioration at which point the pipe collapses (Davies et al. 2001). As a result, it is important to gain a deeper understanding of how sewer defects evolve locally (over time) so as to schedule the appropriate maintenance that will prevent the collapse of a sewer pipe. In summary, the geometrical development of specific defects over the service period of the sewer pipes constitutes a significant research gap in the preventative maintenance of the sewer pipe

## Chapter 3: Methodology

### 3.1 Overview

This research focuses on two main objectives: (1) proposing a framework for calculating the relative growth of defects in underground pipes using image processing techniques (see Figure 3); and (2) validating the framework by implementing it on case studies using real-life sewer pipe images. The process flow underlying in this framework includes three main parts. (1) Image pre-processing is the step-in which images are converted from colour into greyscale form. 2) Feature detection and area scaler (AS) calculation that uses ORB (Oriented FAST (Features from Accelerated Segment Test) and Rotated BRIEF (Binary Robust Independent Elementary Features)) as the feature detection and matching algorithm. 3) Defect segmentation and change calculation. ORB is a rigid and affine invariant algorithm, which means that the algorithm is consistent under any rigid or affine transformation. Rigid deformation includes translation, reflection and rotation. The affine transformation includes, in addition to the rigid transformations, shear and scaling transformations.

The simple form of affine transformation, which covers the rigid transformation and the uniform scaling can be given in Equation (1) as follows (Brown 1992):

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} t_x \\ t_y \end{pmatrix} + s \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \quad (1)$$

where:

$(x_1, y_1)$  coordinate of a point in the first image,  $(x_2, y_2)$  coordinate of a point in the second image,

$\theta$  is the rotation,  $S$  is the uniform scale factor, and  $t$  is the translation

Equation (1) map the point  $(x_1, y_1)$  in the first image into the point  $(x_2, y_2)$  in the second image.

The general 2D affine transformation, which can account for more general spatial distortions such as a shear and changes in aspect ratio, can be expressed as shown in Equation (2):

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix} + \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \quad (2)$$

A shear component of an affine transformation can be given as:

$$Shear_x = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}, \quad Shear_y = \begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix}$$

A shear along one axis causes a distortion of pixels along one axis, proportional to their location on the other axis.

If the scaling in the images is not uniform, the scale along axis  $x$  is different from the scale along axis  $y$  then the aspect ratio is the relative scale between  $x$  and  $y$  axes.

$$Scale = \begin{pmatrix} S_x & 0 \\ 0 & S_y \end{pmatrix}$$

An affine transformation, which can be obtained by applying any combination of translation, rotation, scaling, and shear, describes the cumulative distortions.

ORB detects features in the images, gives the detected feature a description value (descriptor) which depends on the designs and patterns of their neighboring pixel and matches the features. 3) Defect segmentation and change calculation. The common features identified by ORB in the two images are then used to determine the area scaling factor allowing areas encapsulating identified defects to have the same reference. As a result, using the oldest image as a reference, the areas of the defects in newer images will be scaled thus allowing the relative change (in size) for defects of interest to be quantified.

### **3.2 Framework for detection of defect development**

The overview of the proposed system framework is presented in Figure 3.1. Time-series images that are extracted from CCTV inspection videos serve as the original data source. Main processes are developed and implemented in python programming language for which several image processing libraries are readily available such as OpenCV, an open-source library for image and video analysis by Intel. OpenCV is related to computer vision, such as feature and object detection and machine learning(Mistry and Saluja 2016). Other libraries include scikit-image, which is an image processing library that implements algorithms and utilities for use in research, education and industry applications (Walt et al. 2014), and Pillow(PIL), which provides extensive file format support, an efficient internal representation, and fairly powerful image processing capabilities (Clark and Lundh 2020). The main algorithm used is ORB image matching, which has three main components, namely, feature point extraction, generating feature point descriptors, and feature point matching (Luo et al. 2019). Contour, Hough Lines, and curve detection are used for crack segmentation. Some manual processes are required such as extracting frames with defects and drawing triangles bound to the defects in images. The main processes will be discussed in detail in the next sections. The output of the framework is the relative growth of the defect shown in the images in the form of area of the bounding box, approximately bounded rectangle length change, and approximate bounded rectangle width growth.

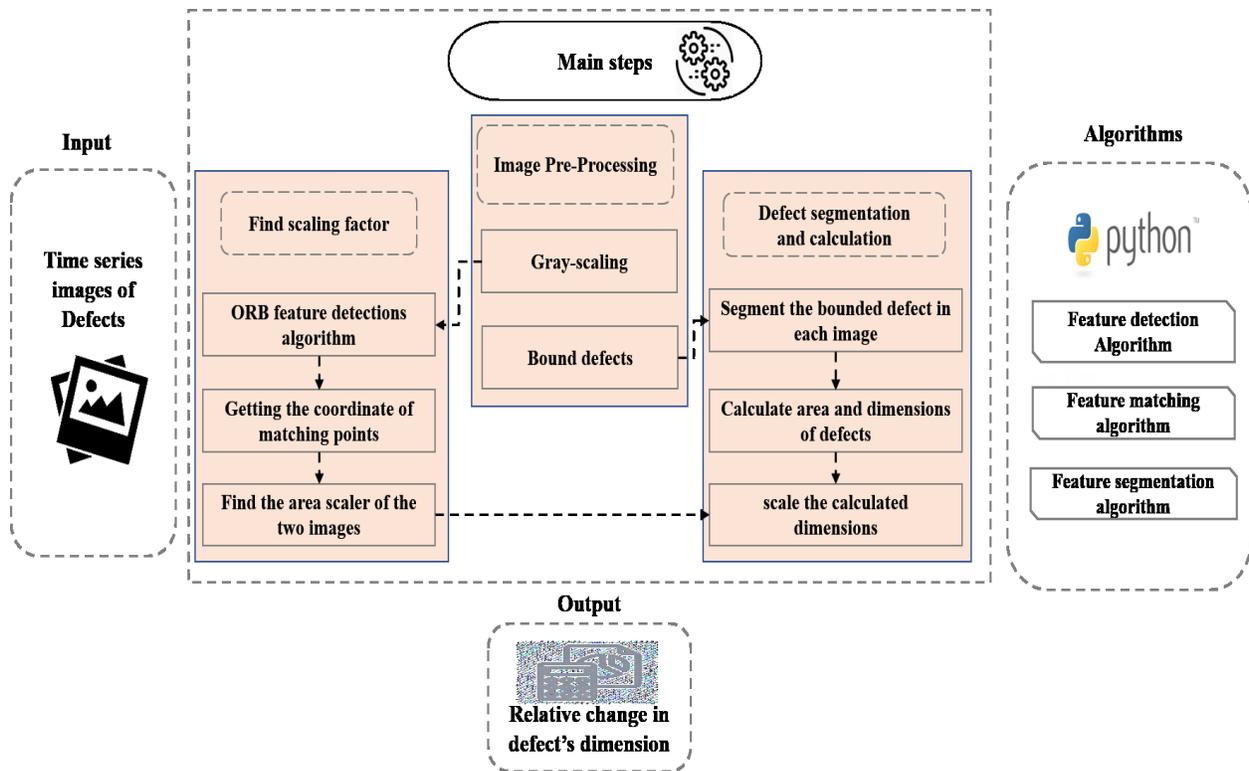


Figure 3. 1: Overview of the automated defect development detection system

### 3.2.1 Input data

Input data of the framework are pairs of images showing the same defect at two different times, as shown in Figure 3.2. These images are extracted from CCTV inspection videos of the same pipe at different times. It is important to mention that when CCTV videos are collected, recordings of the same pipe will not necessarily start at the same point (i.e. from the “zero”) and although the NASSCO standard recommends a speed of 9 m/min (approximately 30 ft/min) for the crawler, in practice this speed cannot be maintained at a constant (Yin et al. 2019). These variabilities during data collection make it impossible to rely on the information reported by the crawler to overlay in a straightforward way the frames from different videos (footage of the same pipe). In this work, a semi-automatic two-step approach is proposed: (1) a deep-learning automatic defect detection and classification procedure is applied to CCTV videos (of the same pipe) to output a list of frames

containing anomalous features (e.g. cracks, fractures, breaks, etc.); and (2) frames representing the same defect are then manually arranged in sets which are then fed to the proposed procedure. An example of the data input is shown in Figure 3.2.



Figure 3. 2 Time-series sewer pipe defect image in 2008 and in 2010

The images show the same scene of pipe segment number 53003 on two different dates: 08/21/2008 and 09/07/2010. After feeding the time-series frames of the same defect to the ORB algorithm in order to calculate AS, the defect of interest should be bounded by drawing a containing rectangle for the defect in each of the images. Luckily, the deep-learning implementation YOLOv3 for image-based feature recognition draws an (approximate) bounding box around each feature of interest it has identified in the image (Yin et al. 2020). As a result, these frames could be used as input to the proposed procedure provided that the bounding boxes are adjusted to best fit the associated defects in order to get an accurate and consistent dimension for the defects. Figure 3.3 shows an example of a defect with a bounding box manually drawn (left) and as reported by YOLOv3 object detection algorithm (right).



Figure 3. 3: Image on the left shows manually bounding the defect using Microsoft PowerPoint software, image on the right shows YOLOv3 defect detection and bounding

### 3.2.2 Output of the framework

When it comes to condition assessment of infrastructure, the introduction of new technologies geared towards improving consistency and efficiency led researchers to explore the benefits of image segmentation and computer vision algorithms to identify and label defects. In this regard, underground drainage pipes are no different, since in recent years CCTV inspections have become very popular worldwide. While traditional inspections of sewers provide valuable information regarding the structural and operational condition of this infrastructure, this information is only a snapshot at a specific time. To extend the capability of the traditional practice of condition assessment, this research explores the possibility of extracting additional value by stacking images in the form of a time series allowing time-dependent data-driven deterioration modelling to be conducted. The outputs of the proposed framework are the relative changes of the defects in the images, which are estimated from the changes in the dimensions of the best fitting bounding-box around the defect of interest. An assumption of the uniform scaling in the images has been made and then a dimension scaler is calculated depending on the calculation of AS. Then, the relative change in the length of the defect is calculated.

### 3.3 Experiments of AS method

#### 3.3.1 Area scaler (AS)

From three matching points that are not aligned in two images, a triangle in each image can be drawn using the coordinates of matching points and the AS can be calculated.

Depending on the calculated  $AS = Area_1/Area_2$ , according to Equation (3), a bounding area of any defect (Area4) in one image can be scaled to be compared with the bounding area of the same defect in the second image (Area3) and the change in the area can be determined as can be shown in Figure 3.4. Thus, the proposed AS could be the tool of correspondence to compare and detect changes between images. However, an accuracy and consistency evaluation should be done to validate the degree of invariance that AS could suffer when taking the effect of deformation result from image acquisition.

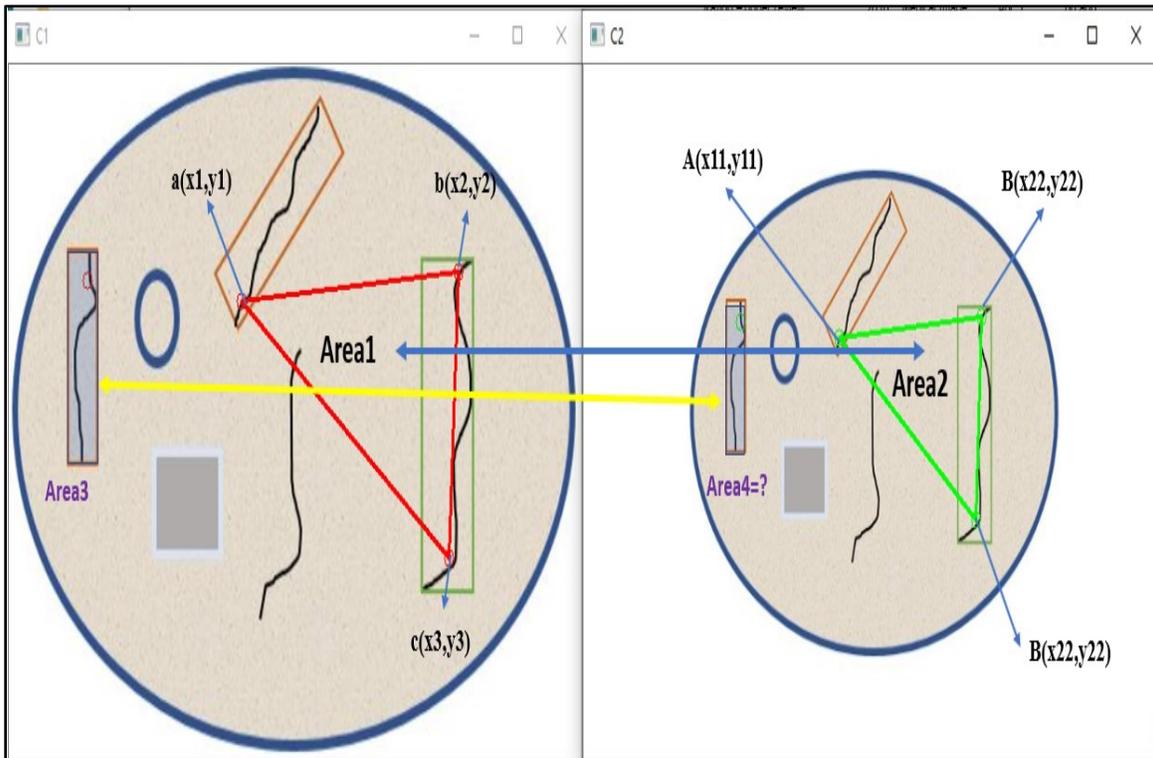


Figure 3. 4: Illustration of the AS

$$AS = Area1/Area2 \quad (3)$$

$$AS = Scale\ Factor^2 \quad \text{when the scaling is uniform} \quad (4)$$

When the coordinates of matching points are known, the area of the formed triangle can be calculated as per:

$$Area_1 = Area_{abc} = \sqrt{S(S - ab)(S - bc)(S - ac)} \quad (5)$$

The lengths of the sides of a triangle can be calculated using Equations (6) through (8):

$$ab = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (6)$$

$$ac = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (7)$$

$$bc = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (8)$$

where  $S$  is the half of the perimeter as calculated using Equation(9):

$$S = \frac{ab + ac + bc}{2} \quad (9)$$

The same calculation is done for the  $Area_2$ .

If AS is consistent (has the same value all over the images), matching areas can be scaled and a growth in an area can be detected.

### 3.3.2 The experiment of AS concept on CCTV images

When the aim of comparing images is measuring changes, a mean of correspondence that can account for variations affecting dimensions of objects in the images is required.

In case of CCTV inspection process, videos are recorded for the same pipe segment at different times. Although these videos are inspecting the same pipe, the geometric characteristic of the video frames could be different. Geometric deformations could be classified as either rigid, affine, or non-rigid deformation. In rigid deformation, the overall relationships between points do not

change. Rigid transformations, which include translation, reflection, and rotation, preserve the points, lines, length, angles, and the parallelism of a set of two lines. This means a triangle in one image maps onto a matching triangle in the second. Affine transformations are restricted to preserve collinearity and allow, besides rotation and translation, also scaling and shearing. Thus, the area scale is constant and does not change all over the image under any affine transformation. Any transformation that falls out of rigid and affine transformation categories is called non-rigid transformation (Davari et al. 2017). An example of a non-rigid transformation, which is popular in the image registration field, is perspective deformation, which is distortion occurring when images of the same scene are taken from different perspectives. Perspective distortion causes imagery to appear smaller the further it is from the camera and more compressed the more it is inclined away from the camera (Brown 1992).

AS is invariant for rigid and affine transformations. Therefore, AS could be a suitable tool to correspond matching images if the purpose is to quantify the change in defect area that occurs between two different times. However, deformations that result from a different perspective of the CCTV camera affect the globality feature of AS. That means when there is a difference in perspective, AS will not be constant all over the image.

To study the effect on the globality of AS due to differences in perspective of the images collected during underground pipe inspections, a set of 49 pairs of sewer pipe matching images taken at different times will be examined.

Data used in this experiment were extracted from two sets of videos collected at different times and provided by our industry partner. Figure 3.5 shows the flowchart of the AS experiment process.

The process will be explained and applied to two pairs of images.

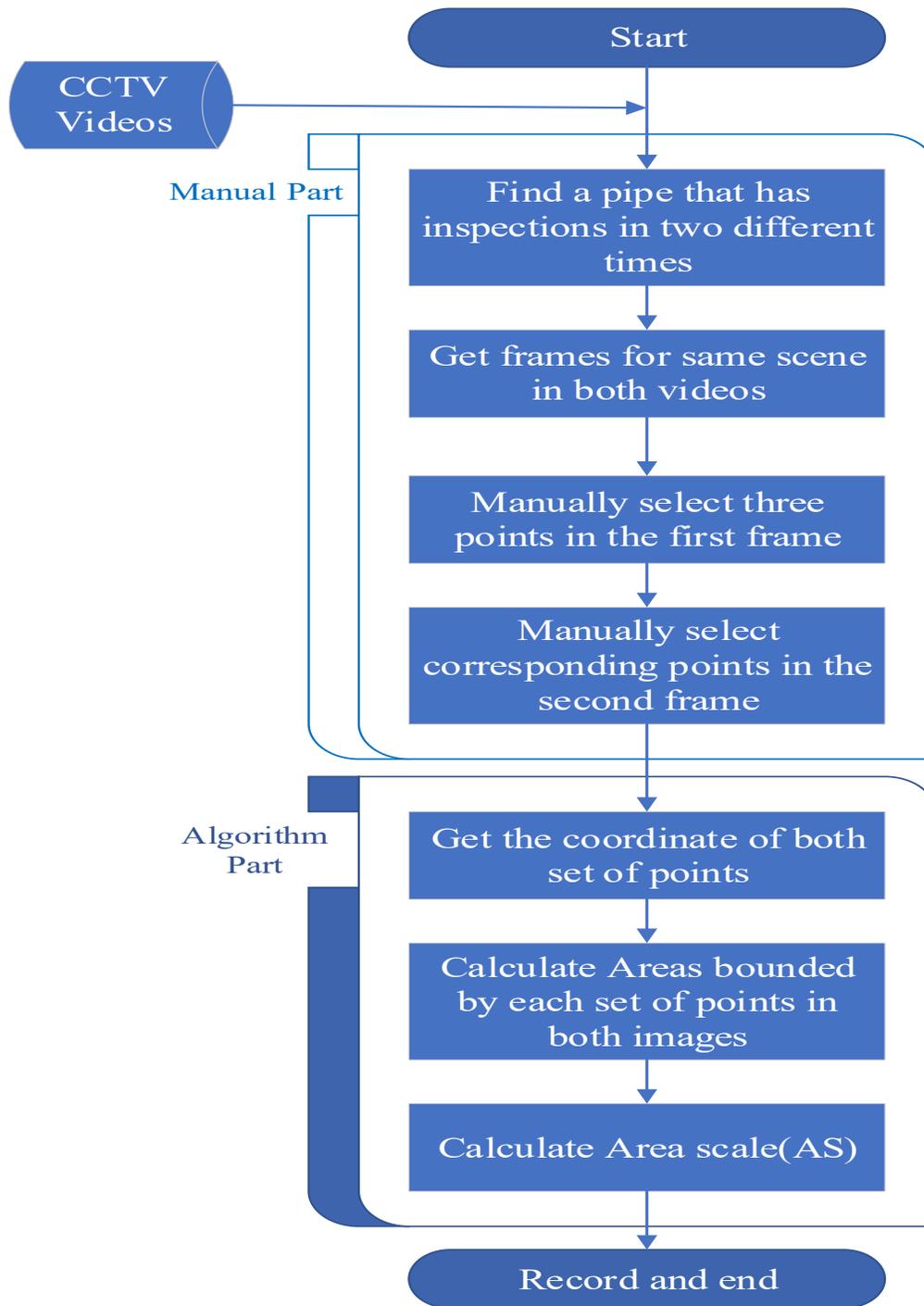


Figure 3. 5: A flowchart of AS experiment process

In the following we show the process applied on two pairs of images:

The images are obtained from inspection videos for two pipe segments which have the IDs 30796 and 53003. For each pair of images, the coordinates of 3 pairs of associated points are shown in

Figures 3.6 and 3.8. To test the globality of AS, the selected points should be distributed all over the images (see Figures 3.7, 3.9). The AS was then calculated, recorded, and is detailed in this section.

### 3.3.2.1 Example 1: Pipe segment: 30796



Figure 3. 6: a) A pair of matching images from pipe segment 30796, b) The first pair of matching areas, c) The Second pair of matching areas, d) The third pair of matching areas

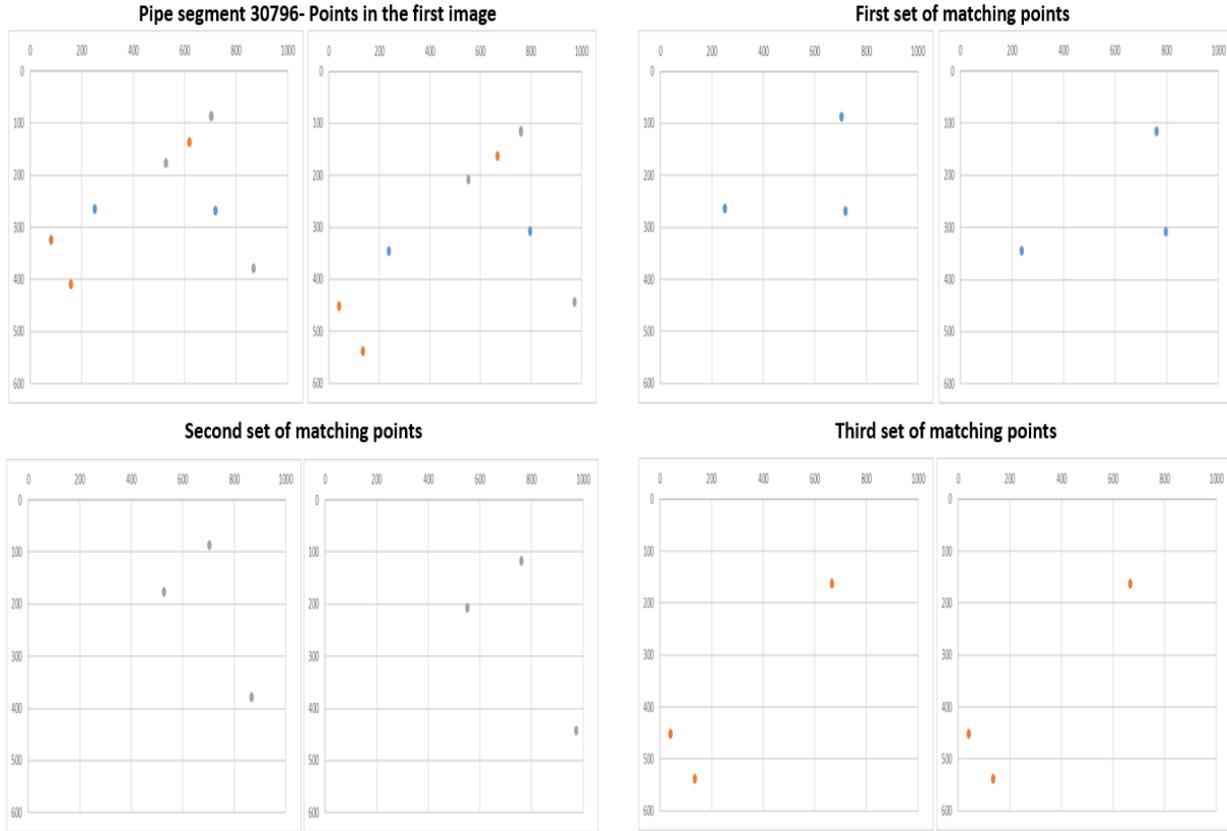


Figure 3. 7: A plot of the selected points in each image.

Using Equations (5) through (9),  $A_1$  and  $A_2$  for each pair of matching areas are calculated, then the AS is calculated using Equation (3).

Table 3. 1: AS calculation for the first pair of images

		Coordinates of matching points						A1	A2	AS = A1/A2
		x1	y1	x2	y2	x3	y3			
Pair1	First image	250	264	704	87	719	268	42414		0.782
	Second image	238	345	760	115	796	307		54252	
Pair2	First image	618	136	158	409	81	324	30060.5		0.737
	Second image	667	163	41	451	135	538		40767	
Pair3	First image	527	177	703	86	868	379	33291.5		0.763
	Second image	552	207	761	117	973	443		43607	

After calculating the three values of AS, further calculations (see Tables 3.2, 3.4) are done to quantify the effect of the perspective deformation on the globality of AS.

Difference-Range represents the error that results if only one set of matching areas was used to calculate AS. It is calculated as Equation (10) as follows:

$$\text{Difference-Range} = |\max(AS_1, AS_2, AS_3) - \min(AS_1, AS_2, AS_3)| \quad (10)$$

Then the average value of AS is calculated using Equation (11).

$$\text{Avg} = (AS_1 + AS_2 + AS_3)/3 \quad (11)$$

Difference-Range (Avg) represents the maximum error that results from using the average of three values of AS and can be determined by the max absolute value of Avg-AS1, Avg-AS2, and Avg-AS3:

$$\text{Difference - Range(Avg)} = \max(|\text{Avg} - AS_1|, |\text{Avg} - AS_2|, |\text{Avg} - AS_3|) \quad (12)$$

In the following, we apply Equations (10) through (12) to calculate the values shown in Table 3.2 for the first pair of images:

$$\text{Difference-Range: } 0.782 - 0.737 = 0.049 = 4.9 \%$$

$$\text{Avg} = (0.782 + 0.737 + 0.763) / 3 = 0.760$$

$$|\text{Avg} - AS1| = |0.760 - 0.782| = 0.021$$

$$|\text{Avg} - AS2| = |0.760 - 0.737| = 0.023$$

$$|\text{Avg} - AS3| = |0.760 - 0.763| = 0.003$$

$$\text{Difference-Range (Avg)} = \max(2.1\%, 2.3\%, 0.3\%) = 2.3\%$$

Table 3. 2: Results summary of AS experiment for the first pair of images

Segment#	Set#	AS1	AS2	AS3	Difference -Range	Average	Avg-AS1	Avg-AS2	Avg-AS3	Difference -range (Avg)
30796	7	0.782	0.737	0.763	4.5%	0.761	0.021	0.024	0.002	2.37%

### 3.3.2.2 Example 2: Pipe segment: 53003

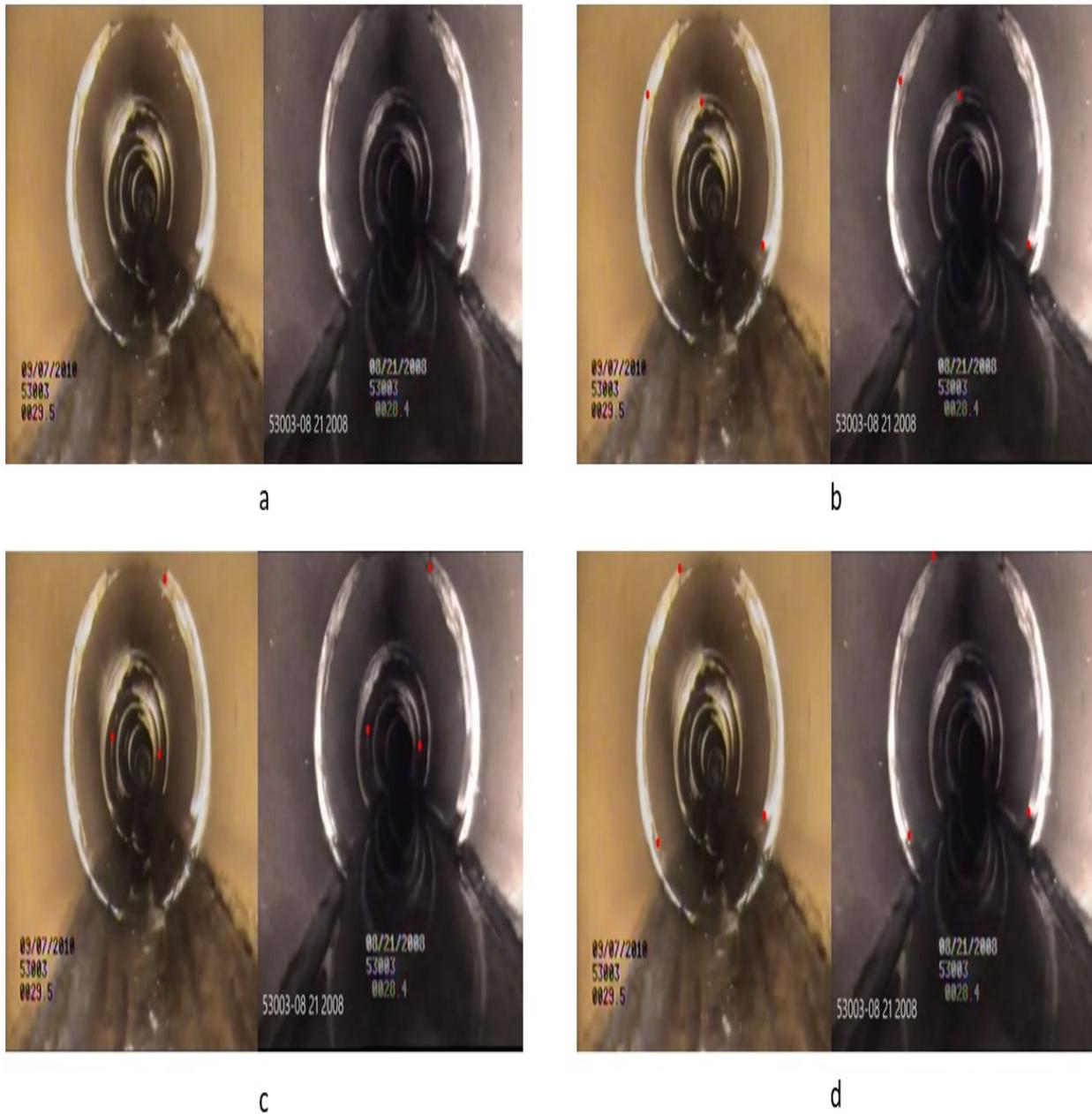


Figure 3. 8: a) A pair of matching images from pipe segment 53003, b) The first pair of matching areas, c) The Second pair of matching areas, d) The Third pair of matching areas

Using Equations (5), (6), (7), (8), (9),  $A_1$  and  $A_2$  for each pair of matching areas are calculated, then the AS is calculated using Equation (3).

Table 3. 3: AS calculation for the first pair of images

Coordinates of matching points		x1	y1	x2	y2	x3	y3	A1	A2	AS = A1/A2
Pair1	First image	341	157	598	169	881	419	30427		0.918
	Second image	317	133	591	152	896	415		33133.5	
Pair2	First image	514	299	765	51	737	327	31166		0.941
	Second image	469	294	731	308	778	290		33111.5	
Pair3	First image	488	30	386	463	885	420	105840.5		0.899
	Second image	463	7	354	449	896	412		117765.5	

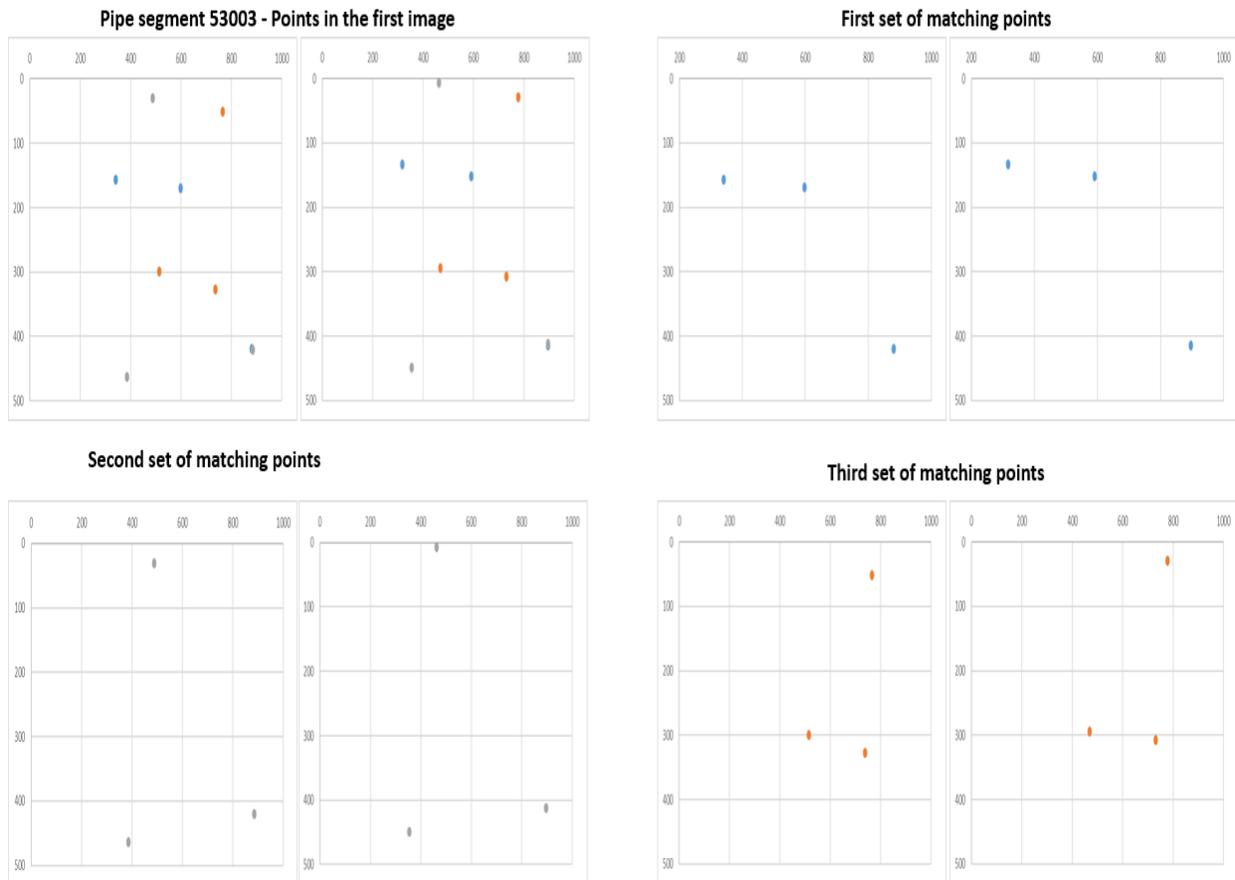


Figure 3. 9: A plot of the selected points in each image.

In the following we apply equations (10, 11, 12) to calculate the values shown in Table 3.4 of the second pair of images:

$$\text{Difference-Range} = 0.941 - 0.899 = 0.042 = 4.2 \%$$

$$\text{Avg} = (0.918 + 0.941 + 0.899)/3 = 0.919$$

$$|\text{Avg} - \text{AS1}| = 0.919 - 0.918 = 0.001$$

$$|\text{Avg} - \text{AS2}| = 0.919 - 0.941 = 0.022$$

$$|\text{Avg} - \text{AS3}| = 0.919 - 0.899 = 0.02$$

$$\text{Difference-Range (Avg)} = \max(0.1\%, 2.2\%, 2\%) = 2.2\%$$

Table 3. 4: Results summary of AS experiment for the second pair of images

Segment#	Set#	AS1	AS2	AS3	Differenc e-Range	Average	Avg-AS1	Avg-AS2	Avg-AS3	Differenc e-range (Avg)
53003	7	0.918	0.941	0.899	4.2%	0.919	0.001	0.022	0.020	2.17%

Table 3.5 shows the result of examining the globality of AS using 49 pairs of images.

Table 3. 5: The results of applying the concept of AS on 49 different pairs of images.

Segment#	Set#	AS1	AS2	AS3	Difference-Range	Avg	Avg-AS1	Avg-AS2	Avg-AS3	Difference-Range (Avg)
30796	1	0.718	0.695	0.729	3.4%	0.714	0.004	0.019	0.015	1.90%
30796	2	0.626	0.655	0.66	3.4%	0.647	0.021	0.008	0.013	2.10%
30796	3	0.765	0.806	0.819	5.4%	0.797	0.032	0.009	0.022	3.17%

30796	4	0.787	0.735	0.724	6.3%	0.749	0.038	0.014	0.025	3.83%
30796	5	0.802	0.797	0.743	5.9%	0.781	0.021	0.016	0.038	3.77%
30796	6	0.633	0.66	0.617	4.3%	0.637	0.004	0.023	0.020	2.33%
30796	7	0.782	0.737	0.763	4.5%	0.761	0.021	0.024	0.002	2.37%
30796	8	0.613	0.627	0.633	2.0%	0.624	0.011	0.003	0.009	1.13%
53003	1	0.741	0.719	0.771	5.2%	0.744	0.003	0.025	0.027	2.73%
53003	2	0.886	0.893	0.862	3.1%	0.880	0.006	0.013	0.018	1.83%
53003	3	0.903	0.885	0.93	4.5%	0.906	0.003	0.021	0.024	2.40%
53003	4	0.995	1	1.013	1.8%	1.003	0.008	0.003	0.010	1.03%
53003	5	0.791	0.764	0.783	2.7%	0.779	0.012	0.015	0.004	1.53%
53003	6	0.678	0.714	0.699	3.6%	0.697	0.019	0.017	0.002	1.90%
53003	7	0.918	0.941	0.899	4.2%	0.919	0.001	0.022	0.020	2.17%
53003	8	0.713	0.696	0.693	2.0%	0.701	0.012	0.005	0.008	1.23%
53003	9	0.96	0.909	0.911	5.1%	0.927	0.033	0.018	0.016	3.33%
53003	10	1.028	0.986	1.001	4.2%	1.005	0.023	0.019	0.004	2.30%
53003	11	0.953	0.954	0.935	1.9%	0.947	0.006	0.007	0.012	1.23%
53003	12	0.899	0.92	0.901	2.1%	0.907	0.008	0.013	0.006	1.33%
53003	13	0.91	0.922	0.934	2.4%	0.922	0.012	0.000	0.012	1.20%
53003	14	0.957	0.999	0.984	4.2%	0.980	0.023	0.019	0.004	2.30%
53003	15	0.984	1.012	1.032	4.8%	1.009	0.025	0.003	0.023	2.53%
53003	16	0.837	0.876	0.828	4.8%	0.847	0.010	0.029	0.019	2.90%
53003	17	0.861	0.818	0.851	4.3%	0.843	0.018	0.025	0.008	2.53%
53003	18	0.855	0.874	0.858	1.9%	0.862	0.007	0.012	0.004	1.17%

52759	1	0.879	0.899	0.918	3.9%	0.899	0.020	0.000	0.019	1.97%
53052	1	0.857	0.899	0.876	4.2%	0.877	0.020	0.022	0.001	2.17%
53052	2	0.866	0.874	0.821	5.3%	0.854	0.012	0.020	0.033	3.27%
53052	3	0.964	0.979	0.931	4.8%	0.958	0.006	0.021	0.027	2.70%
85707	1	0.875	0.898	0.906	3.1%	0.893	0.018	0.005	0.013	1.80%
85707	2	0.741	0.763	0.751	2.2%	0.752	0.011	0.011	0.001	1.13%
85707	3	0.839	0.821	0.795	4.4%	0.818	0.021	0.003	0.023	2.33%
85707	4	0.601	0.593	0.583	1.8%	0.592	0.009	0.001	0.009	0.93%
85707	5	0.978	0.99	0.963	2.7%	0.977	0.001	0.013	0.014	1.40%
85707	6	0.8	0.835	0.845	4.5%	0.827	0.027	0.008	0.018	2.67%
85707	7	0.825	0.826	0.807	1.9%	0.819	0.006	0.007	0.012	1.23%
85707	8	0.679	0.712	0.697	3.3%	0.696	0.017	0.016	0.001	1.70%
85707	9	0.918	0.912	0.89	2.8%	0.907	0.011	0.005	0.017	1.67%
85729	1	0.917	0.922	0.907	1.5%	0.915	0.002	0.007	0.008	0.83%
85729	2	0.892	0.894	0.924	3.2%	0.903	0.011	0.009	0.021	2.07%
85729	3	0.8	0.767	0.774	3.3%	0.780	0.020	0.013	0.006	1.97%
85729	4	0.713	0.763	0.737	5.0%	0.738	0.025	0.025	0.001	2.53%
85729	5	0.893	0.941	0.937	4.8%	0.924	0.031	0.017	0.013	3.07%
85729	6	0.962	0.976	0.929	4.7%	0.956	0.006	0.020	0.027	2.67%
85729	7	0.754	0.768	0.799	4.5%	0.774	0.020	0.006	0.025	2.53%
85729	8	1.069	1.067	1.068	0.2%	1.068	0.001	0.001	0.000	0.10%
85729	9	0.902	0.916	0.925	2.3%	0.914	0.012	0.002	0.011	1.23%
85729	10	1.016	1.082	1.085	6.9%	1.061	0.045	0.021	0.024	4.50%

The results show that the maximum recorded variation in AS value over a pair of images is 6.9%, (see Figure 3.10), which can be related to the effect of camera angle variation. This means that calculating the AS value using only one set of matching points will give an overall accuracy of approximately 93%. Thus, finding an area of a defect could have a maximum inaccuracy of 6.9%.

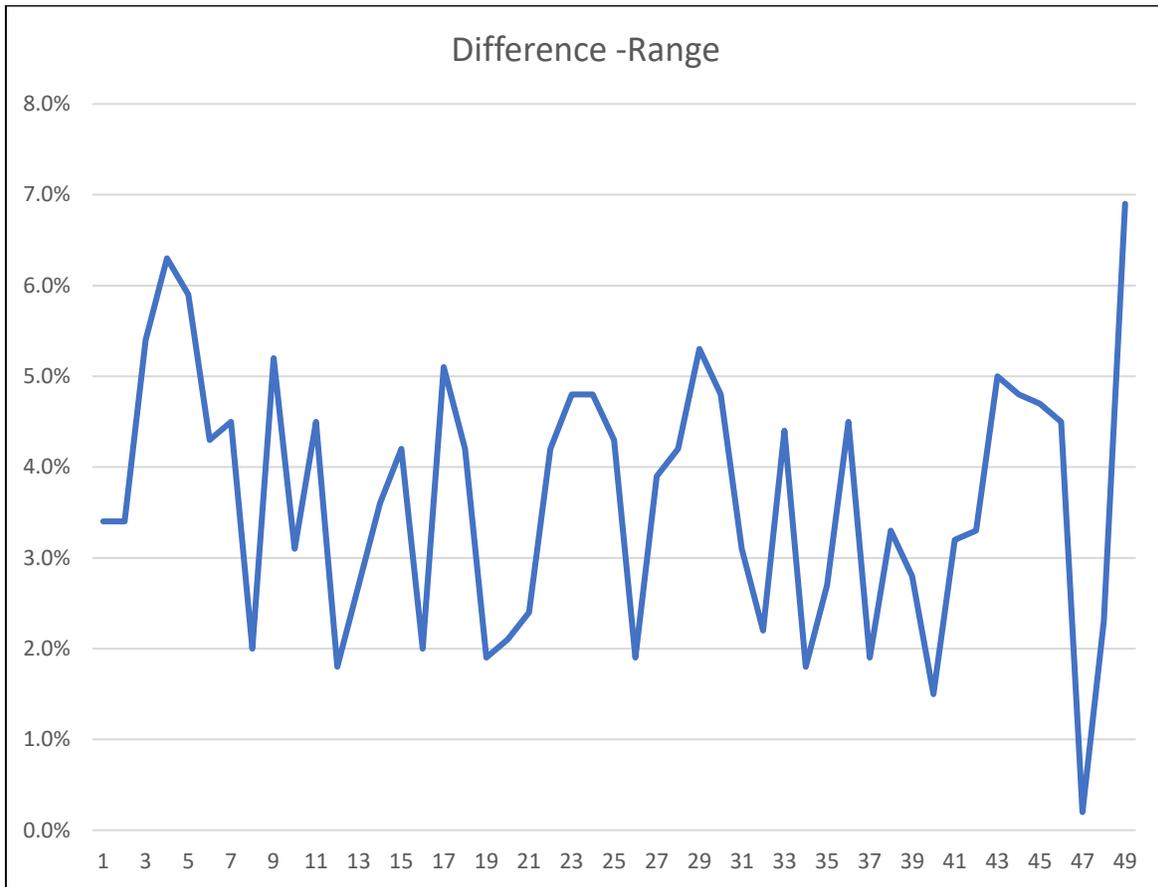


Figure 3. 10: Maximum recorded variation in AS value over a pair of images

However, taking the average AS value of three sets of matching points in a pair of matching images would improve the overall accuracy of the scaling process. As the results demonstrate, when taking the average AS of three sets of matching points, the maximum variation recorded over all the pairs is 4.5% and this error is within reasonable bounds for practical estimates (see Figure 3.11).

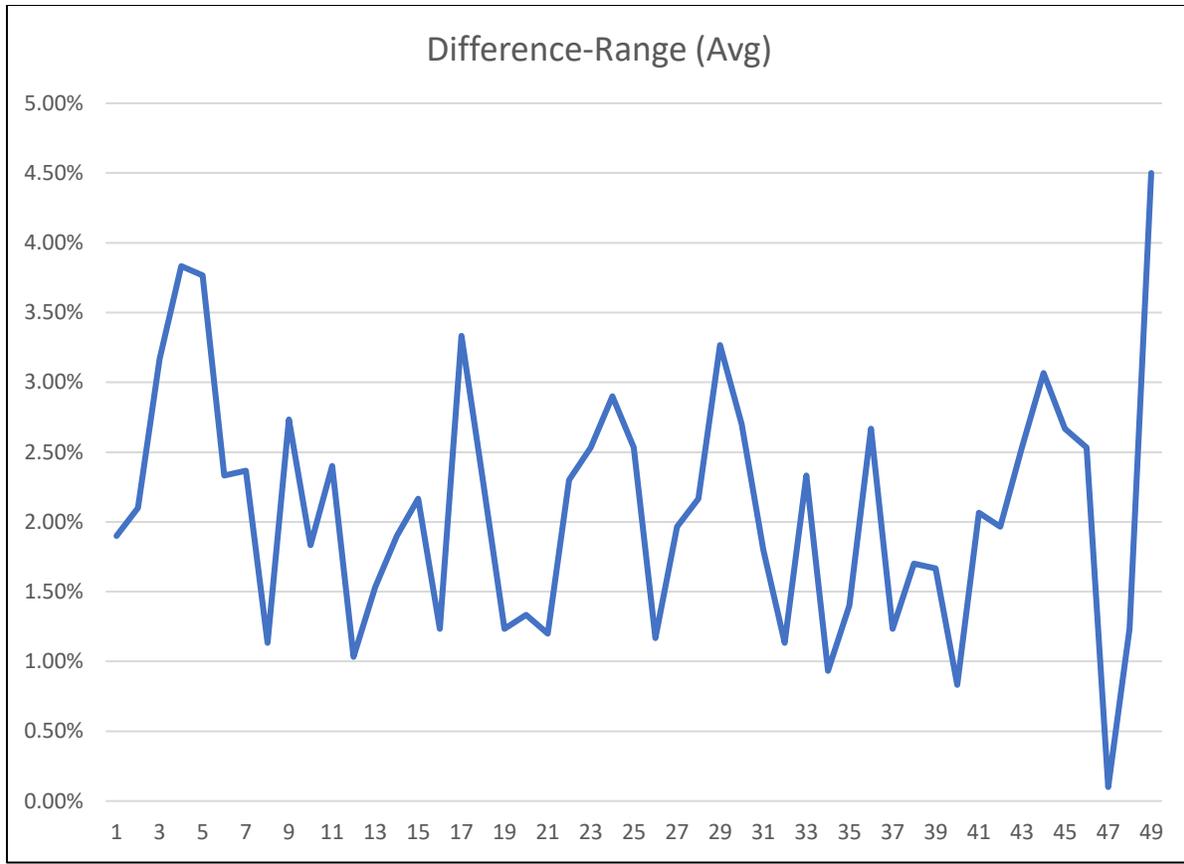


Figure 3. 11: Maximum variation in AS taking the average of three iterations

To summarize the results, Figure 3.12 shows the comparison between the Difference-Range and the Difference-Range (avg) from our 49sets of experiments. It can be shown from the figure that the value of Difference-Range (Avg) is lower than the Difference-Range in terms of the overall trend (see the line chart in Figure 3.12), mean value and deviations (see the boxplot and histograms in Figure 3.12), The lower number of the indicator means that using the average value of 3 iterations (Avg) is more stable than taking a single value of AS. In conclusion, taking the average improves the accuracy of AS to reach 95.5 %.

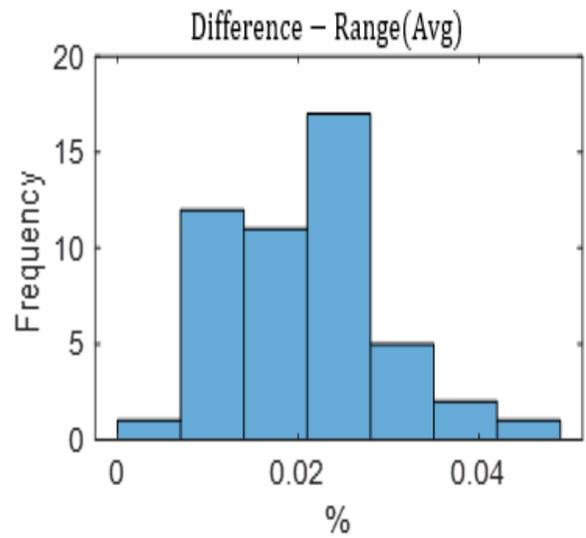
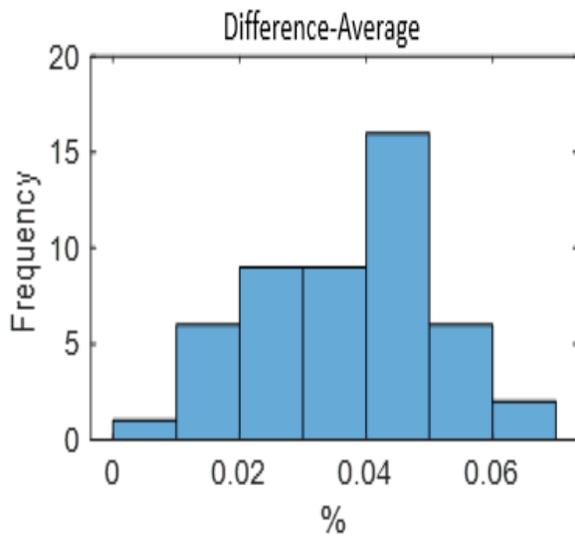
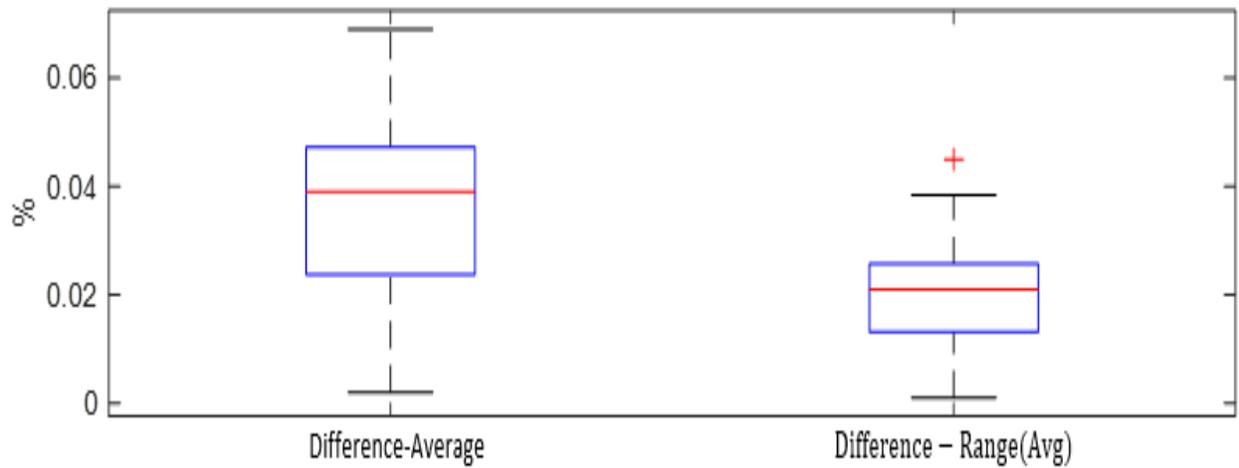
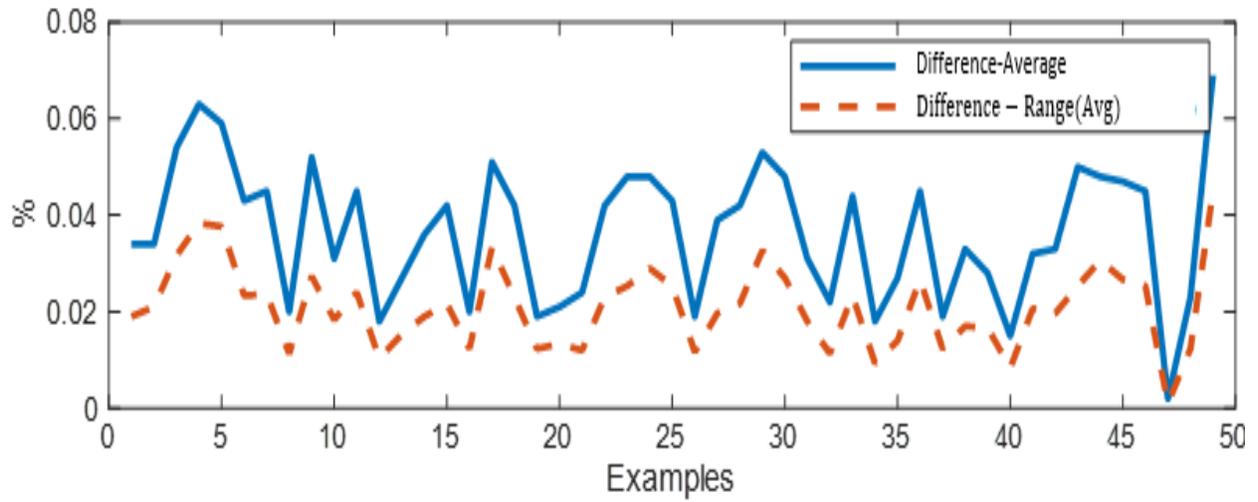


Figure 3. 12: Comparison between “Difference-Range” and “Difference-Range (Avg)”

## **Chapter 4: Implementation Process of the Framework**

The implementation process can be divided into three sections, namely, 1) image pre-processing, 2) feature detection and AS calculation, and 3) defect segmentation and change calculation. The three boxes (i.e., those shown in red, blue, and yellow) in Figure 4.1 indicate what is included in each of the three sections. The red box includes image preparation and pre-processing, which is described in Section 4.1. Pairs of images that have been prepared in the red box processes will be fed to the algorithm in the blue box and the yellow box. The blue box includes the application of the feature detection algorithm, feature matching algorithm, and the AS calculation, which altogether will be described in more detail in Section 4.2. The yellow box includes the defect segmentation and relative change calculation, which will be explained in Section 4.3.

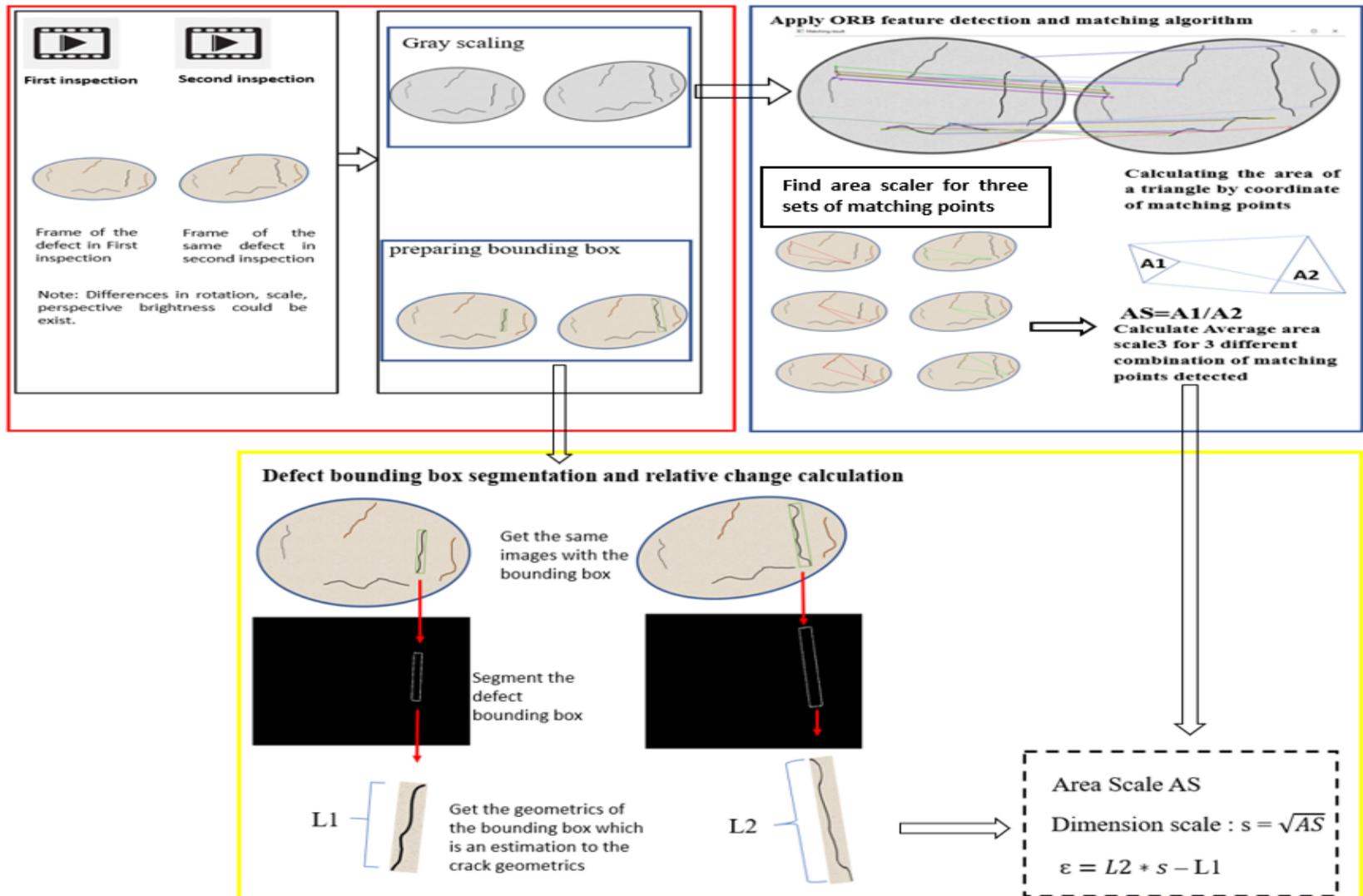


Figure 4. 1: Implementation of the framework proce

## 4.1 Preparation and preprocessing

### 4.1.1 Data preparation

In this step, the original pairs of images without bounding boxes are converted to grayscale to be processed by the feature-matching algorithm in order to get AS. The defects in the images are delimited by manually adjusting the boxes determined by the deep-learning defect detection algorithm. Given that the bounding boxes are proxies for the evolution of the defects, they need to fit (perfectly) the features of interest. The images with the bounding boxes will be fed to the segmentation algorithm in order to calculate the dimensions of the bounding box and estimate the relative change of the defect. Figure 4.2 shows a flowchart to properly prepare the input images for feeding the framework.

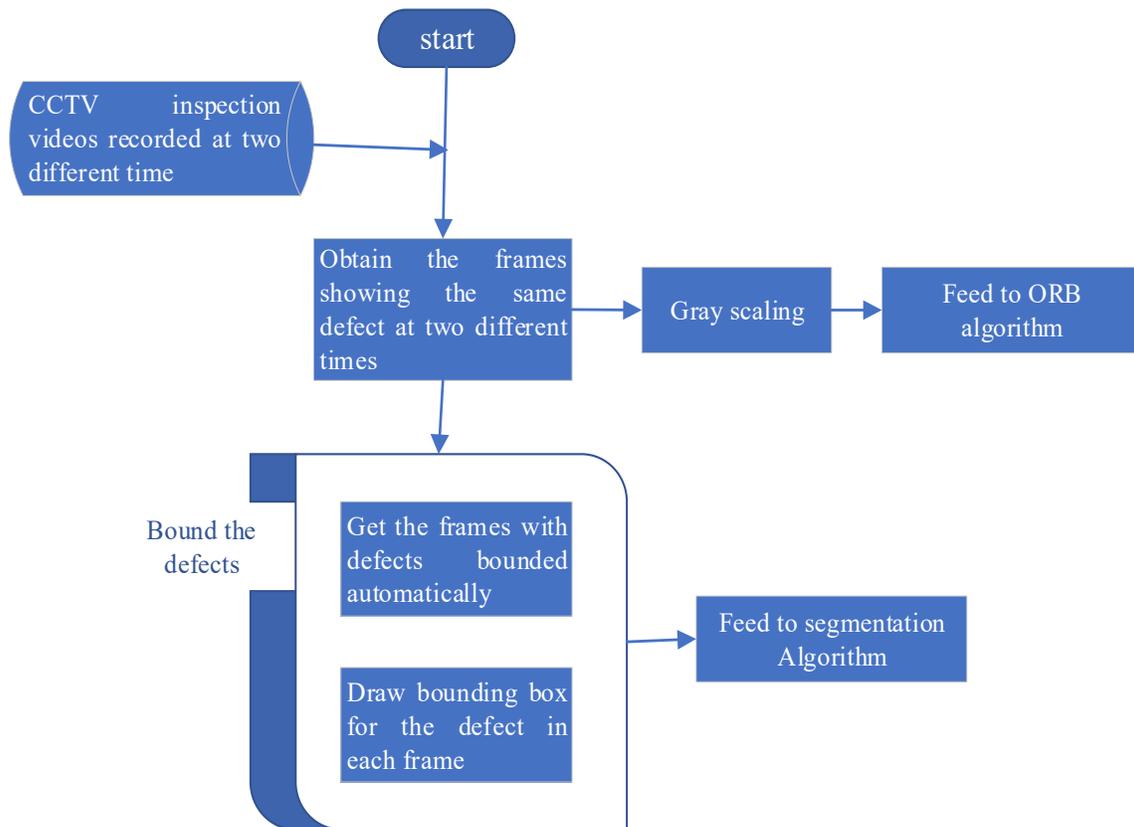


Figure 4. 2: Flow chart of input data preparation

### 4.1.2 Grey scale conversion

Converting the color image into a greyscale image is one of the more important and technically useful image processing techniques used for various purposes (Saravanan 2010). For some image processing applications, color images have no additional merit than greyscale images. In fact, it requires a complex algorithm to deal with color images and increased the computational requirements. A greyscale image has one channel of color which is the grey, where a colour image has three channels, which are red, green and blue. Greyscale images contain brightness information only and no colour information (Padmavathi and Thangadurai 2016) There are many different ways to do the colour to greyscale conversion, and, in the following section, three of them are discussed. The basic and most simple method is the arithmetical average method by taking the average of the three colours in the colour image (Kanan and Cottrell 2012), as shown in Equation (13):

$$Greyscale = (R + G + B)/3 \quad (13)$$

This method is very simple, but it does not account for the effect of the color wavelengths and how humans perceive them.

The lightness method averages the most prominent and least prominent colors, as shown in Equation (14):

$$Greyscale = (\max(R, G, B) + \min(R, G, B)) / 2 \quad (14)$$

The weighted method or luminosity method considers how the human eye perceives the colour and accounts for the lighting information. Since the colour red has more wavelength of all the three colours, and green is the colour that not only has less wavelength than red, but also green is the

colour that has a more soothing effect on the eyes (Kanan and Cottrell 2012). The luminosity conversion is given by Equation (15):

$$\text{Greyscale} = ((0.299 * R) + (0.587 * G) + (0.114 * B)) \quad (15)$$

Equation (15) gives more weight, i.e., 0.587, to the colour that contributes most, which is the green, and decreases the contribution of the red to 0.299 and the blue to 0.114.

Figure 4.3 shows the application of the three previously explained methods of greyscaling on a sample image. Although it is difficult to discern the difference between the different methods, Kanan and Cottrell (2012) concluded that luminosity method helps in reducing the loss of information carried by the human brightness perception and in improving the performance of feature detection algorithms, especially when there is significant change in illumination conditions, which is one main noise factor in underground pipe inspection images.

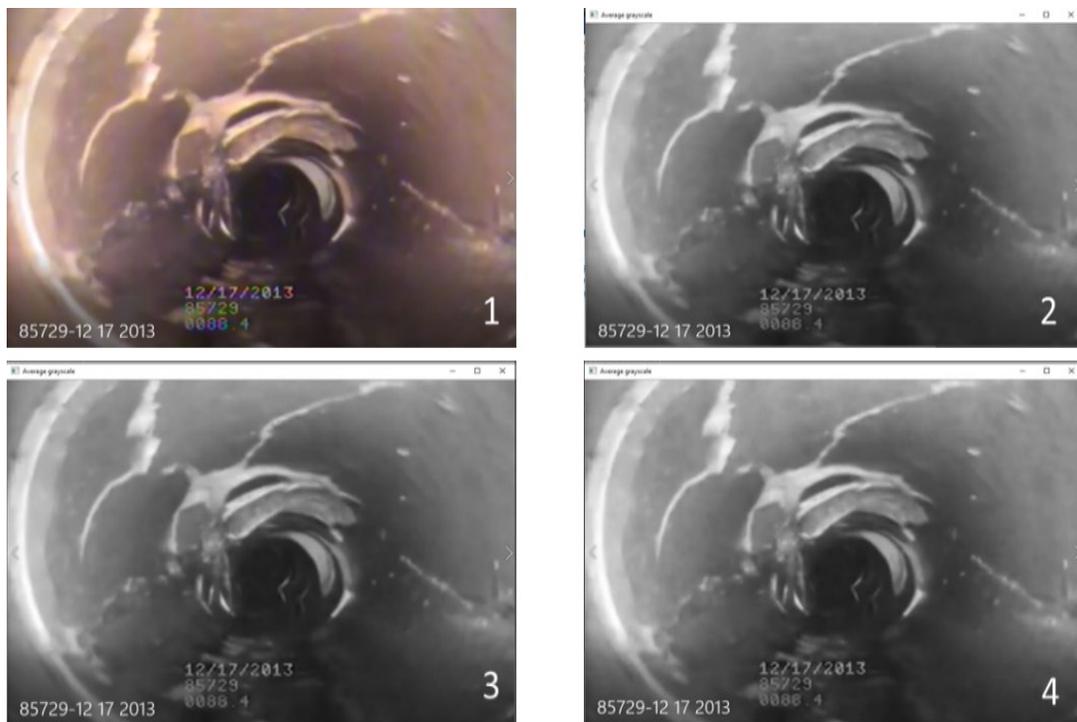


Figure 4. 3: Applying different grey scaling methods, 1) coloured image, 2) average greyscaling method, 3) lightness greyscaling method, 4) luminance greyscaling method

The ORB matching algorithm, which will be discussed in the following section, operates on greyscale images. It compares the intensity of one pixel with others. The intensity of a pixel is its grey colour degree that varies in the range between 0 which is black and 255 which is white. The ORB feature adopts the third method, the luminosity method, to make the greyscale conversion.

#### **4.2 ORB feature detection and matching:**

The first step of image registration is feature detection and description. In order to detect points of interest and key points in an image, or what are generally called features, a feature detector is needed. Feature detectors work on identifying features which are typically corners, lines, edges, intersections etc. (Tareen et al. 2018). When features are detected, they are described based on the designs and patterns of their surrounding pixels. This process is significant as it is considered the foundation of later matching detected features. Detected features are assigned with a distinctive value that is used for the purpose of recognition and matching. ORB is an affine invariant feature-detector (Tareen et al. 2018).

FAST (Features from Accelerated Segment Test) detector covers the feature detection process in the ORB Algorithm. FAST has been improved by adding an orientation component (Rublee et al. 2011). The well-known algorithm known as the binary robust independent elementary features (BRIEF) algorithm performs feature matching (Calonder et al. 2010). BRIEF is a bit string description of an image patch constructed from a set of binary intensity tests. After the description step is done, feature matching can be performed using one of many available methods such as Threshold based matching; Nearest Neighbor; Nearest Neighbor Distance Ratio etc. However, as BRIEF is a binary descriptor, Hamming distance is the feature-matching method that is used with BRIEF (Tareen et al. 2018). Figure 4.4 shows a flowchart that illustrates the process of ORB algorithm.

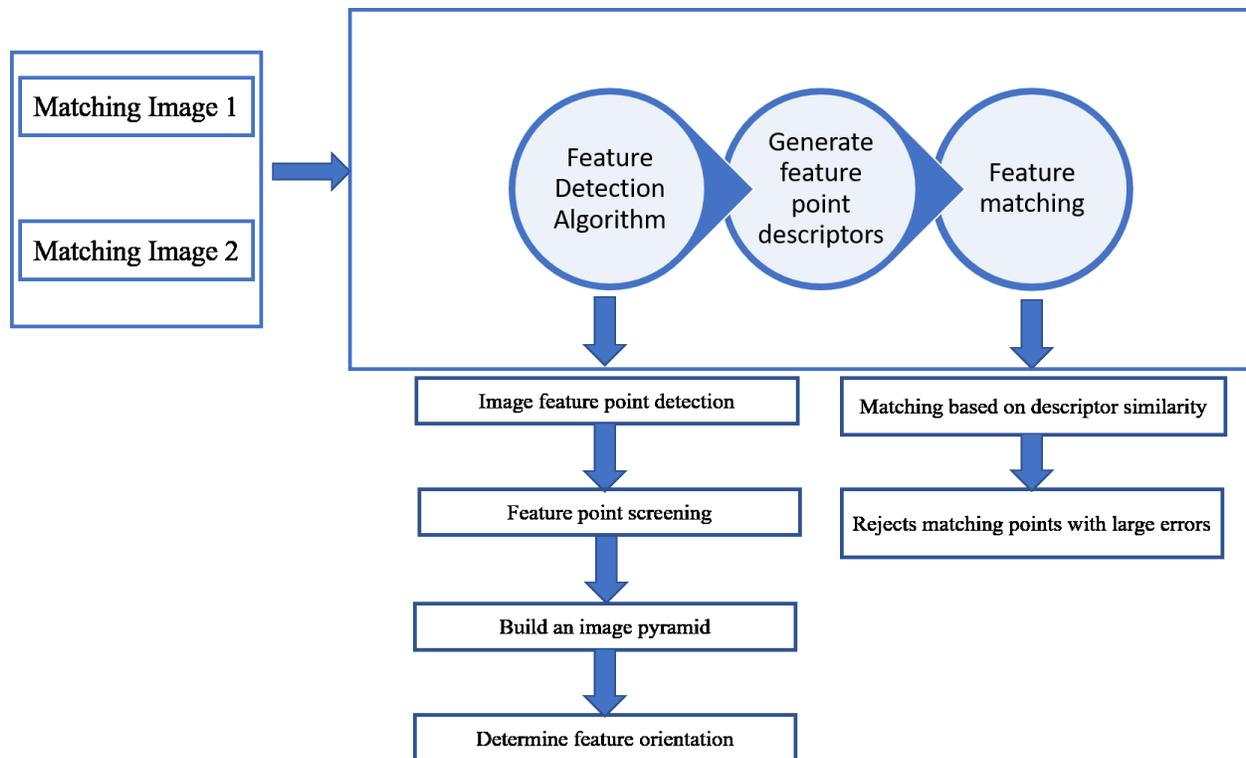


Figure 4. 4: Flow chart of image matching process based on the ORB algorithm (Luo et al. 2019)

#### 4.2.1 FAST corner detection algorithm

Many corner detection algorithms have been developed and used for various purposes, such as tracking, localization, image matching, and recognition. Most corner detection algorithms depend on computing corner response (C), which is a score function to determine if a point is a candidate to be a corner by comparing the response score to a threshold value. If the response value of the candidate corner is greater than the specified threshold value, then a further step is performed to ensure that only one corner is retained in a specific spot(local maximum) (Rosten and Drummond 2006).

##### 4.2.1.1 Basic FAST key-points detection

FAST uses the intensity of the candidate pixel in order to determine if it is a corner or not. In simple words, a pixel is a corner candidate if there is a substantial difference in intensity between

that pixel and the neighboring pixels (Luo et al. 2019). Rosten and Drummond (2005) introduced the corner response test as follows: considering a circle of 16 pixels around the corner candidate which is denoted with  $p$ . Then,  $p$  is considered a corner if  $n$  adjacent pixels of the surrounding circle have more intensity than  $p$  by a specified threshold  $t$  or  $n$  adjacent pixels of the surrounding circle have less intensity than  $p$  by a specified threshold  $t$ . In other words,  $n$  adjacent pixels of the surrounding circle are all brighter than  $p$  by a threshold  $t$  or  $n$  adjacent pixels of the surrounding circle are all darker than  $p$  by a threshold  $t$ .

The following image shown in Figure 4.5, which is from a study by Rosten and Drummond (2006), illustrates the test of corner detection with a circle of 16 pixels.

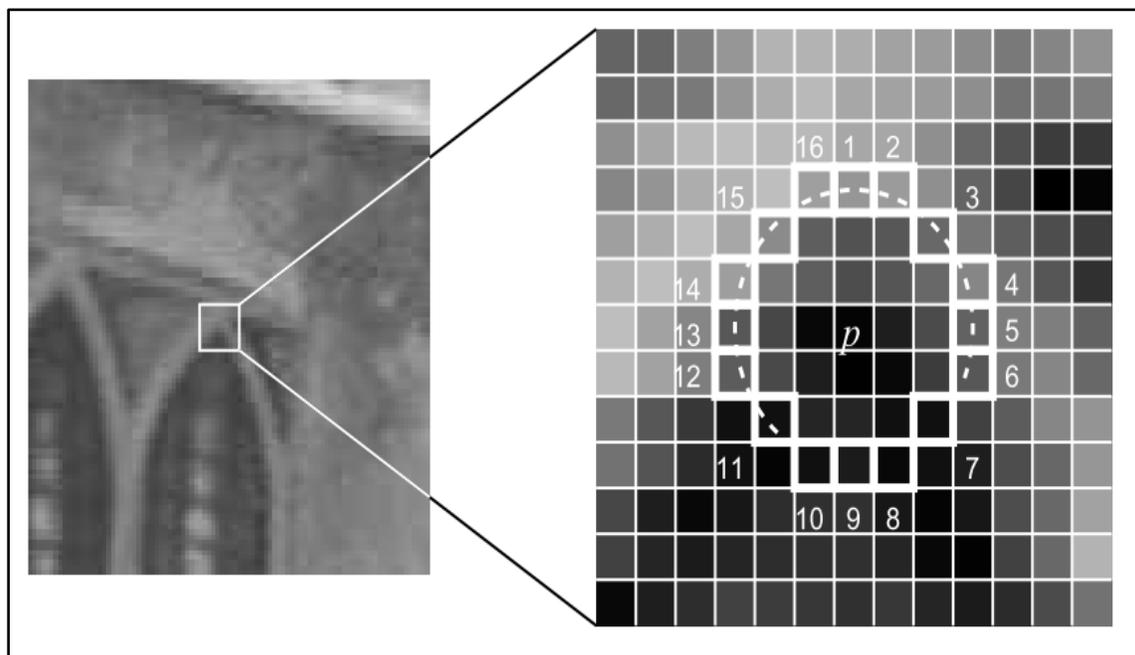


Figure 4. 5: Corner detection test Rosten and Drummond (2006)

In the image,  $n$  was given as 12,  $p$  is classified as a corner as there is at least  $n = 12$  continuous pixels represented by the dotted line that are brighter than  $I_p + t$ .

In order to find the local maximum, a score function to compare the confidence of a candidate to be a corner is required. The candidate with the highest score is considered a corner and the neighboring candidates with a lower score are removed. FAST uses the sum of the absolute difference between the pixels in the contiguous arc and the center pixel.

$$V = \max \left( \sum_{x \in S_{bright}} |I_{p \rightarrow x} - I_p| - t, \sum_{x \in S_{dark}} |I_p - I_{p \rightarrow x}| - t \right) \quad (16)$$

where:

$V$  is the sum of the absolute intensity difference between the pixels in the contiguous arc and the center pixel;

$I_p$  is the intensity of the center corner;

$I_{p \rightarrow x}$  is the intensity of a surrounding pixel; and

$t$  is the intensity threshold.

#### 4.2.1.2 Feature point screening

ORB adds an improvement to the basic FAST detector by applying the Harris function to calculate the response and sort the corners detected by the basic FAST (Harris and Stephens 1988(Harris and Stephens 1988; Luo et al. 2019)Luo et al. 2019). The Harris function is calculated as follows in Equations (17) and (18):

$$R = \text{Det}(M) - K * (\text{Tr}(M))^2 \quad (17)$$

$$M = \sum W(x, y) \begin{bmatrix} I_x^2 & I_x * I_y \\ I_x * I_y & I_y^2 \end{bmatrix} \quad (18)$$

where:

$R$  is the Harris score value;

$M$  is a  $2 \times 2$  matrix;

$K$  ranges from 0.04 to 0.06;

$W(x, y)$  is image window

$I_x$  is the variation of the feature point in the horizontal direction; and

$I_y$  is the variation of the feature point in the vertical direction.

After calculating the Harris response score for the detected corner, they are sorted and then  $N$  points with the highest score are selected to the next step of matching.

#### **4.2.1.3 Image scale pyramids**

The image pyramid gives a flexible, convenient multiscale resolution presentation that reflects the multiple scales of processing in the human visual system (Choudhary et al. 2012). To improve the scale invariance of FAST, getting the same result even if the size of the image changes, Rublee et al. (2011) proposed the construction of image scale pyramids, which is a multiscale representation of the image at different resolutions. These pyramids are scale-invariant in the sense that an object's scale change is offset by shifting its level in the pyramid. Intuitively, this property enables a detector to detect objects across a large range of scales by scanning the model over both positions and pyramid levels (Zhao et al. 2019). Figure 4.6 shows an example of a multiscale pyramid of an image. Each scaling level of the pyramid is examined by applying FAST algorithm to detect the key-points in the image with respect to both the spatial coordinates and the scale level in the pyramid. Thus, image features at different scales are located effectively and the scale invariance characteristic is improved in ORB feature detection and matching algorithm.

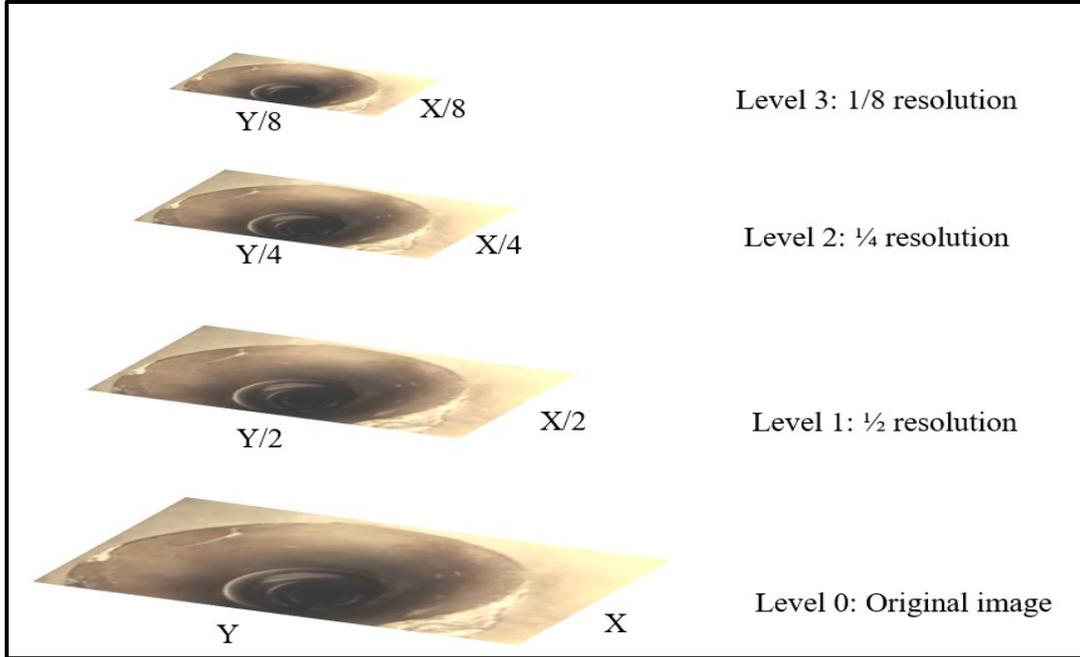


Figure 4. 6: A pyramid or multiscale representation of a single image

#### 4.2.1.4 Feature points orientation

The orientation of detected features can be described in different methods. For example, SIFT algorithm depends on histograms of gradient computations and SURF Algorithm uses approximation by block patterns. The basic FAST algorithm does not have an orientation operator, therefore another improvement to basic FAST was added by (Rublee et al. 2011) which is an orientation depending on the concept of intensity centroid. “The intensity centroid assumes that a corner's intensity is offset from its center, and this vector may be used to impute an orientation”. The following shows the computation of intensity centroid (Luo et al. 2019; Rublee et al. 2011; Rosin 1999).

The moment of the image block is calculated as Equation (19):

$$m_{pq} = \sum_{x,y \in B} x^p y^q I(x, y) \quad p, q = \{0,1\} \quad (19)$$

where  $x$  and  $y$  are pixel coordinates, and  $I(x, y)$  is the grey value of the corresponding pixel. Thereafter, the centroid is calculated depending on the moment according to Equation (20):

$$C = \left( \frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \quad (20)$$

where  $m_{00}$  calculated for  $p = 0$  and  $q = 0$  is referred to as the mass of the image block while  $m_{01}$  and  $m_{10}$  are known as the centroid of the image block.

The final step is to get the orientation of the block by obtaining the direction vector  $\overrightarrow{OC}$  according to Equation (21):

$$\theta = \arctan(m_{01}/m_{10}) \quad (21)$$

These added orientation calculations make the algorithm more accurate by improving the rotation invariance.

#### **4.2.2 BRIEF (binary robust independent elementary features)**

The main idea of the BRIEF matching algorithm is to simply compare the intensities of pairs of points surrounding the candidate feature to create a binary descriptor (Calonder et al. 2010). It relies on a relatively small number of intensity difference tests to represent an image patch as a binary string. This approach was inspired by earlier work on classifying image patches depending on a small number of pairwise intensity comparisons that are used as training data to classification trees or naive Bayes classifiers. On the other hand, when there are a relatively large number of detected key points, the dimensionality of the data will be relatively high. Thus, relying on the Hamming distance between the key points in two matched images is inefficient (Gionis et al. 1999). However, the binary structure of the descriptors facilitates the use of different searching algorithms to make the matching process faster and more efficient.

After applying the FAST detection algorithm to find the key-points in an image, the BRIEF algorithm is applied to create a binary feature descriptor that contains only values of 1s and 0s. The length of this binary vector falls in the range from 128 to 512 bits string.

The first step in BRIEF is to smooth the image in order to reduce the negative effect of high-frequency noise on the descriptor, thus increasing the stability and repeatability of the descriptor. For the purpose of smoothing, Gaussian kernel is used. The next step is to sample random pairs of pixels in the patch of the key-point, which is a pre-defined square of selected width and height pixels around the key-point, considering the size of patch  $p$  is  $(S \times S)$  and assuming the key-point is located in the center of the patch. Calonder et al. (2010) tested five different distributions to sample from:

1.  $(x, y) \sim i. i. d. Uniform(-S/2, S/2)$

$x$  and  $y$  pixels in the random pair is sampled using uniform distribution  $(-S/2, S/2)$  around the key-point;

2.  $(x, y) \sim i. i. d. Gaussian(0, \frac{1}{25} S^2)$

Gaussian distribution with  $\mu = 0$  and  $\sigma = \frac{1}{5} S$  around keypoint is used to sample  $x$  and  $y$  pixels in the random pair.

3.  $x \sim i. i. d. Gaussian\left(0, \frac{1}{25} S^2\right), y \sim i. i. d. Gaussian(x, \frac{1}{100} S^2)$

A Gaussian distribution centred around the key-point with mean  $\mu = 0$  and a standard deviation  $\sigma = \frac{1}{5} S$  is used to sample the first pixel. The second pixel ( $y$ ) in the random pair is drawn from a Gaussian distribution centred around the first pixel ( $x$ ) with mean  $\mu = xi$

and a standard deviation of  $0.1 \times S$ . This forces the test(pair) to be more local. Test(pair) locations outside the patch are clamped to the edge of the patch.

4. *The  $(x_i, y_i)$  are randomly sampled from discrete locations of a coarse polar grid introducing a spatial quantization.*

Both x and y pixels in the random pair are sampled from discrete locations of a coarse polar grid introducing a spatial quantization.

5.  *$\forall i : x_i = (0,0)$  and  $y_i$  is takes all possible values on a coarse polar grid containing  $n_d$  points.*

The first pixel(x) in the random pair is at (0, 0) and the second pixel(y) in the random pair is sampled from discrete locations of a coarse polar grid

Sampling method number 2 has been found to have relatively preferable advantages over other distributions, thus it has been accredited to be the sampling method with ORB algorithm. In the ORB algorithm, distribution number 2 is used to perform the sampling which is a gaussian distribution centred around the key-point. BRIEF selects the number of N random pairs, tests them, and assigns a value of 1 if the first pixel is brighter than the second pixel, and a value of 0 otherwise.

The test is defined on patch P of size  $S \times S$  as:

$$\tau(P; x, y) := \begin{cases} 1 & \text{if } P(x) < P(y) \\ 0 & \text{other wise} \end{cases} \quad (22)$$

where  $P(x)$  is the pixel intensity or the grey value at the first pixel in the randomly N pairs chosen, and  $P(y)$  is the grey value at the second pixel in that pair (Calonder et al. 2010; Luo et al. 2019).

After testing the intensity values of each of the N random sampled pair pixels, an N dimensional vector with N binary string can be assigned as a descriptor for the given key-point using the formula shown in Equation (23):

$$f_N(p) = \sum_{1 \leq i \leq N} 2^{i-1} \tau(P; x_i, y_i) \quad (23)$$

#### 4.2.2.1 rBRIEF: Rotation-aware BRIEF

Since BRIEF is not designed to be rotationally invariant, which means that the descriptor calculated for a patch around a key-point will change if the patch is subject to a rotation (Calonder et al. 2010). Thus, its performance falls off dramatically when the image is rotated. In other words, BRIEF does not have a rotation invariance and its rotation tolerance is no more than a few degrees. Rublee et al. (2011) improved the BRIEF descriptor by adding orientation information calculated at the key-point, thus, steering BRIEF according to key-point orientation (Rublee et al. 2011). A rotation matrix  $R_\theta$  is obtained as shown in Equation (24):

$$R_\theta = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \quad (24)$$

For each detected key-point,  $N$  tests are performed to obtain the binary descriptor. Thus a  $2 \times n$  matrix defines the location  $x, y$  of the tests as follows:

$$S = \begin{pmatrix} x_1 \dots \dots, x_N \\ y_1 \dots \dots, y_N \end{pmatrix}$$

Then the steered version of  $S$ ,  $S_\theta$  is obtained using the patch orientation  $\theta$  and the corresponding rotation matrix  $R_\theta$ :

$$S_\theta = R_\theta \times S \quad (25)$$

Finally, the steered directional descriptor can be obtained as shown in Equation (26):

$$g_n(P, \theta) := f_N(p) | (x_i, y_i) \in S_\theta \quad (26)$$

#### 4.2.3 Feature point matching

The last step in the feature matching process is to decide if two key points are matches. That is done by measuring the similarity between two features that are detected at two images taken at different times. As BRIEF algorithm describes detected features as a binary vector of 0,1 forms,

the Hamming method is used to measure the distance between two given binary vectors that describes the two features. The Hamming distance between two binary vectors is the number of bits that must be changed in one vector to get the second one. It is the number of places where two vectors  $u$  and  $v$  differ. Where  $u, v$  are the vectors that have been obtained from BRIEF descriptor for the given features. The simple Hamming distance measure is efficient when the dimensionality of the data is small; however, in an image matching process, the dimensionality of the data could be relatively large so a more effective search structure for performing similarity search over high dimensional data has garnered much interest (Gionis et al. 1999). In rBRIEF, locality sensitivity hashing (LHS) is used as the nearest neighbour searcher. Hash tables are used to store the descriptors of the key points. Then the matching buckets for a given descriptor are determined and a brute force matching technique is applied to compare the given descriptor with all the elements of the matching buckets.

### **4.3 Defect segmentation and change calculation**

The last technique to be applied in the proposed framework is extracting the bounding box of the defect in order to calculate its development. The extracting process depends on isolating the range of the colour that the bounding box has. For example, if the bounding box is green, we will try to extract the green range by defining the lower value of components (G, R, B) of green colour and the upper values. Once we segment the bounding box, the Hough line technique is applied to find coordinates of the bounding rectangle corners.

Once the coordinates of the bounding box are ready, the bounding box area is calculated. Thereafter, the change in the length of the bounding box, which is an estimation of the defect length is calculated.

#### 4.4 Case studies

##### 4.4.1 Case study 1

In order to validate the framework, it will be applied to real images of defects in sewer pipes. The images are extracted from CCTV inspection videos that were taken in 2012 and 2013. The first pair of images are for a crack (see Figure 4.7). Table 4.1, 4.2, and 4.3 shows the calculation of the area scaler values in the first case study.

Table 4. 1: The calculation of the first value of Area Scaler in the first case study

Matching Points coordinates	C1		C2	
	X	Y	X	Y
	172.8	59.04	198	60.48
	184.8	132.6	209	129
	89.85	124.41	114.04	129.6
Area	3069.68		3256.13	
Area Scaler	0.94			

Table 4. 2: The calculation of the second value of Area Scaler in the first case study

Matching Points coordinates	C1		C2	
	X	Y	X	Y
	172.8	59.04	198.72	60.48
	184.8	123.6	209	129
	102.02	99.53	124.41	104.5
Area	2527.69		2771.96	
Area Scaler	0.92			

Table 4. 3: The calculation of the third value of Area Scaler in the first case study

Matching Points coordinates	C1		C2	
	X	Y	X	Y
	184.8	123.6	209	129
	102	99.53	124.4	104.5
	93.6	131.04	121.2	139.2
Area	30396		32773	
Area Scaler	0.93			

$AS_1 = 0.94$  First value of AS

$AS_2 = 0.92$  Second value AS

$AS_3 = 0.93$  Third value of AS

$$Average AS = \frac{0.94 + 0.92 + 0.93}{3} = 0.93$$

Difference-Range (Avg)=  $\max(1\%, 0\%, 1\%) = 1\%$  the maximum error that can be result in this case study.

Assuming that the image is scaled uniformly, the scaling factor is given as following:

$$S = \sqrt{AS} = \sqrt{0.93} = 0.964$$

An estimation of the relative change in the length of the defect will be calculated. The length of the best fit rectangle of the crack in the first image is 117 pixels and in the second image is 127 pixels, then the relative change is given as following:

$$\varepsilon = L_2 \times s - L_1 = 127 \times 0.964 - 117 = 4.92 \text{ Pixel}$$

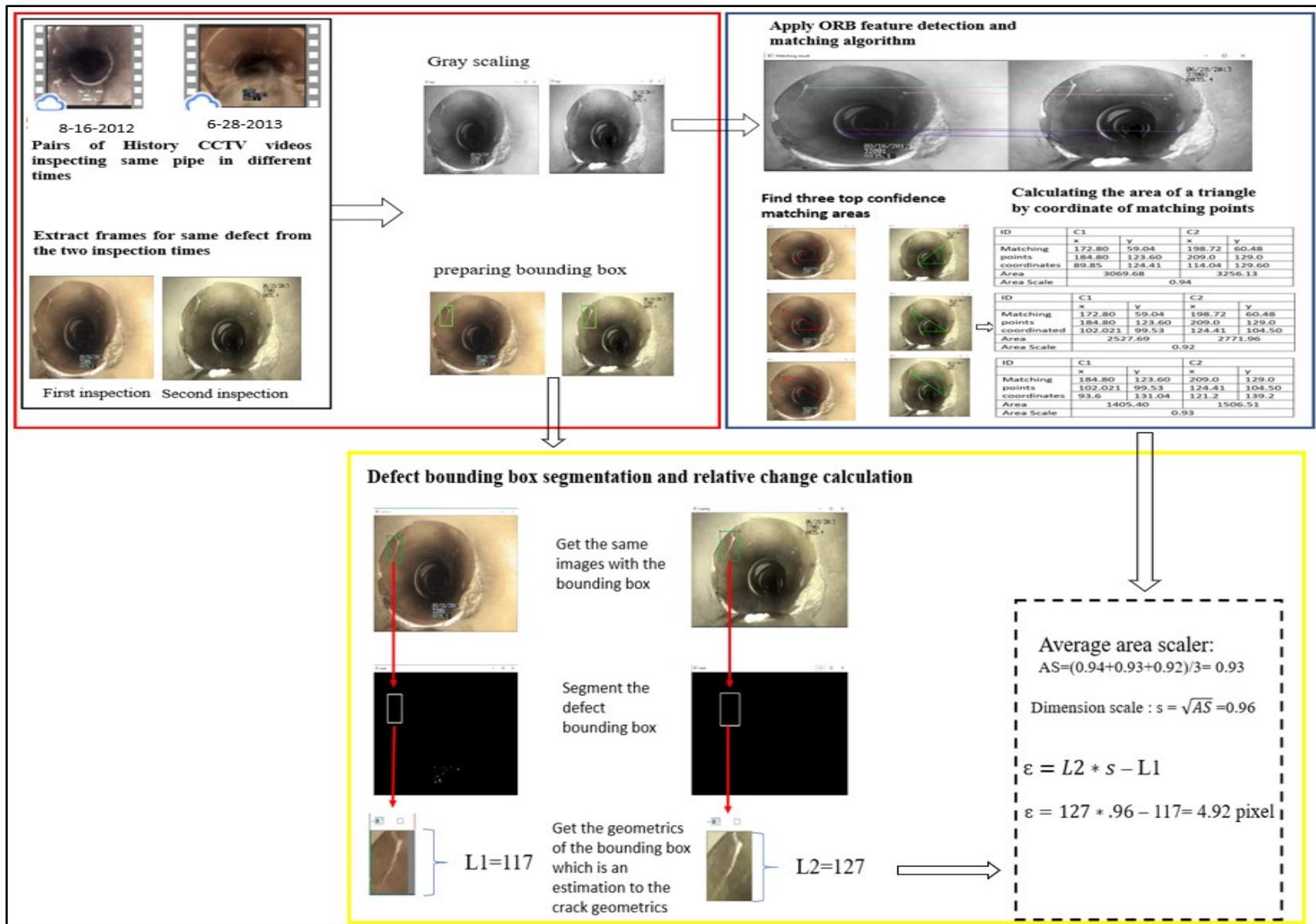


Figure 4. 7: Implementation of the framework for the first case study

#### 4.4.2 Case study 2

The proposed framework is applied on a second pair of images for a crack extracted from CCTV inspection videos that were taken in 2008 and 2010. Table 4.4, 4.5, and 4.6 shows the calculation of the area scaler values in the first case study. Figure 4.8 shows the implementation of the framework on the images of the crack.

Table 4. 4: The calculation of the first value of Area Scaler in the second case study

Matching Points coordinates	C1		C2	
	X	Y	X	Y
	414	238	380	231
	643	319	630	307
	1067	390	1100	373
Area	9042		9606	
Area Scaler	0.94			

Table 4. 5: The calculation of the second value of Area Scaler in the second case study

Matching Points coordinates	C1		C2	
	X	Y	X	Y
	437	547	411	525
	573	144	563	119
	1105	382	1133	367
Area	123382		134558	
Area Scaler	0.916			

Table 4. 6: The calculation of the third value of Area Scaler in the second case study

Matching Points coordinates	C1		C2	
	X	Y	X	Y
	392	523	359	503
	653	200	637	181
	426	248	392	229
Area	30396		32773	
Area Scaler	0.921			

$AS_1 = 0.94$  First value of AS

$AS_2 = 0.916$  Second value AS

$AS_3 = 0.921$  Third value of AS

$$\text{Average AS} = \frac{0.94 + 0.916 + 0.921}{3} = 0.926$$

Difference-Range (Avg)=  $\max(0.5\%, 1\%, 1.4\%) = 1.4\%$  the maximum error that can result in this case study.

Assuming that the image is scaled uniformly, the scaling factor is given as follows:

$$S = \sqrt{AS} = \sqrt{0.926} = 0.962$$

An estimation of the relative change in the length of the defect will be calculated. The length of the best fit rectangle of the crack in the first image is 194 pixels and in the second image is 201 pixels, then the relative change is given as follows:

$$\varepsilon = L_2 \times s - L_1 = 201 \times 0.962 - 194 = 0.638 \text{ Pixels}$$

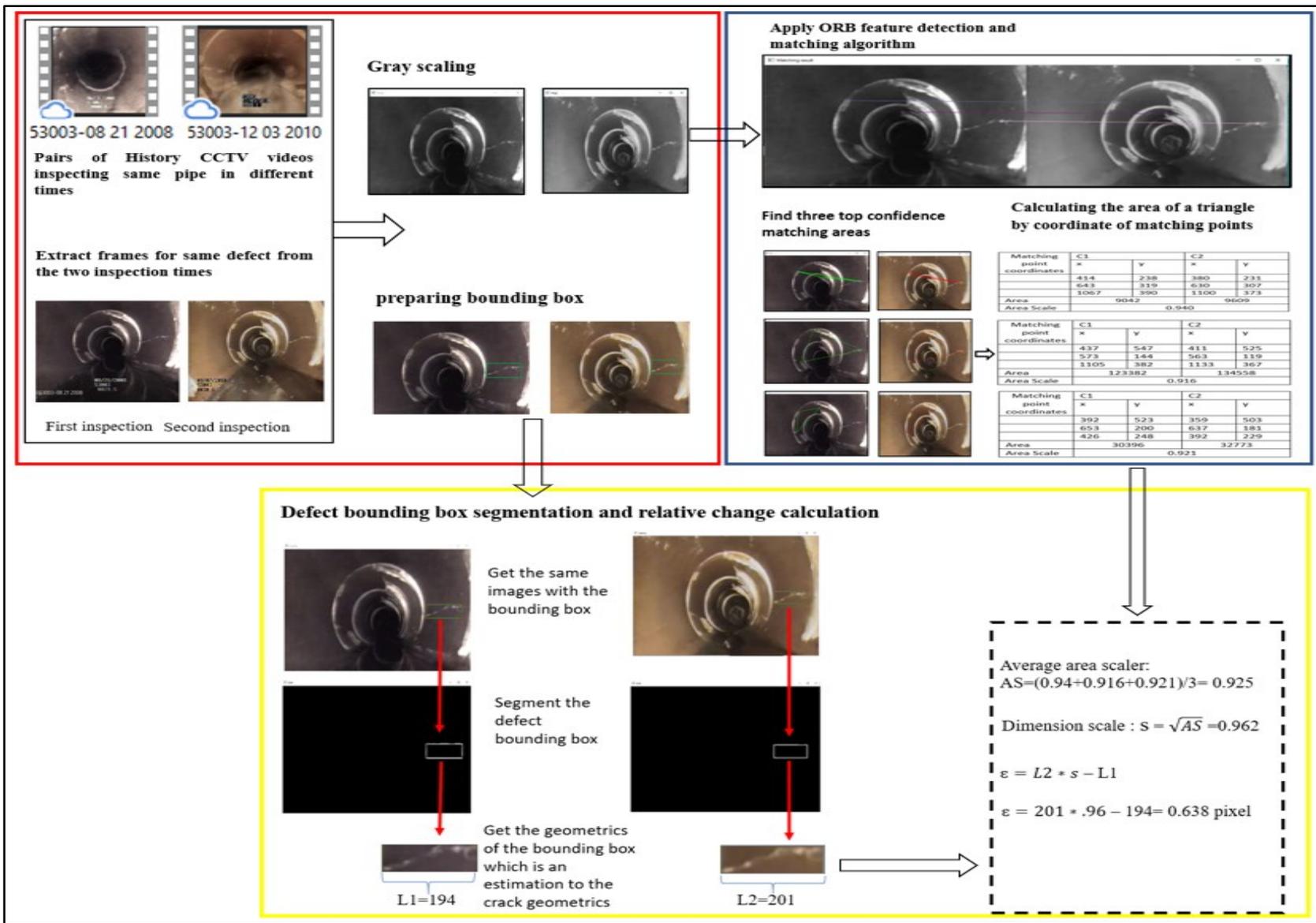


Figure 4. 8: Implementation of the framework in the second case study

The validation of the framework is limited to visual validation as there is no available measurement of the defects. For the first case study, visually, a small growth in the crack could be noticed. The result reflects the visual inspection and a small increase in the crack length is measured. The growth is estimated to be 4.92 pixels.

For the second case study, the crack seems to be unchanged and no development can be detected visually. When applying the framework, it would be assumed the framework should give no change in the crack length. The calculated measurement of the defect development is 0.638 pixels.

The result of the framework is a relative change in pixels and future researches can be focused on how to integrate the information of the camera specification and the geometric distance between the camera and the defect in order to convert this measurement into metric units.

## **Chapter 5: Conclusion and future work**

### **5.1. General conclusion**

With the advent of new technologies in the field of image segmentation and computer vision algorithms, both the efficiency and the accuracy have been greatly improved and that brings many benefits in terms of identifying defects and defect labelling. In this regard, CCTV inspections of underground drainage pipes have become very popular worldwide. Although conventional inspections of sewers provide valuable information about the structural and operational condition of this infrastructure, this information is discrete and only a snapshot at a specific time. However, the ability to access and interpret the information inherent in a time series of images, such as the development of defect, is still a gap in this field of research. In this research, image processing algorithms have been employed to extract the relative change in defect growth from images obtained from CCTV inspection videos. However, the change in the dimension of a defect is measured in pixels, thus, it is still a relative change. The AS factor has proven to be an effective means to correspond matching images. However, the results show that the maximum recorded variation in AS value over a pair of images is 6.9% which can be related to the effect of camera angle variation. This means that calculating the AS value using only one set of matching points will give a maximum error of approximately 6.9%.

However, taking the average AS value of three sets of matching points that represent areas from different parts of the set of matching images improves the overall accuracy of the scaling process. As the results demonstrate, when taking the average AS of three sets of matching points, the maximum error that could be resulted is 4.5%.

## **5.2. Research contribution**

The research extends the capability of the traditional practice of condition assessment, this contribution explores the possibility of extracting additional value by stacking images in the form of a time series allowing time-dependent data-driven deterioration modelling to be conducted. AS factor has been experimented with as a corresponding mean. It has proven to be an effective means to correspond matching images with an accuracy of 95.5%. The ORB detection and matching algorithm was integrated with other pieces of python codes to implement and automate the main process of the introduced framework.

The outputs of the proposed framework are the relative changes of the defects in the images, which are estimated from the changes in the dimensions of the bounding-box best fitting the defect of interest. An assumption of the uniform scaling in the images has been made and then a dimension scaler is calculated depending on the calculation of AS. Then, the relative change in the length of the defect is calculated.

## **5.3. Research limitations**

- The performance of the framework depends on the quality of the images, noisy and low-quality images will affect the framework at the feature matching step giving false matching features.
- The calculated output of the framework depends on the bounded rectangular boxes on the defect. The object detection algorithm sometimes gives inaccurate bounding boxes, which may affect the results.
- The calculation of the dimension scale factor neglects the scaling inconsistency over the image. Images with a large difference between vertical scale factor and horizontal scale factor can negatively affect the accuracy of the output of the framework. However, this

error could be avoided by taking the frames with the original size that normally have uniform scaling factors horizontally and vertically.

#### **5.4.Recommendation and future work**

The developed framework can be a foundation for future studies on deterioration modelling and condition assessment of underground pipes, such as the following:

- The framework has the potential to be developed and integrated with a defect detection algorithm such as YOLO, which employed in the field of underground pipes by Yin et al. (2020)Yin et al. (2020) to automatically provide the framework input data.
- The ORB algorithm was chosen in this research in the role of feature matching algorithm; however, there are other algorithms that can be used in order to determine the algorithm that results in the best algorithm.
- In this research, three different values were used to calculate the average value of AS because of the limited number of correct matching points. As higher quality images can be obtained, increasing the number of values to calculate the average AS will be possible and will result in higher accuracy.
- In this research, the concept of area scaler is used as a corresponding measure between the given images; however, there is the possibility to experiment with more complicated transformation models in order to further improve the accuracy of the framework.

## References

- Adhikari, R. S., O. Moselhi, and A. Bagchi. 2014. "Image-Based Retrieval of Concrete Crack Properties for Bridge Inspection." *Automation in Construction* 39: 180–94.  
<https://doi.org/10.1016/j.autcon.2013.06.011>.
- Allouche, E., and P. Freure. 2002. "Management and Maintenance Practices of Storm and Sanitary Sewers in Canadian Municipalities." *Institute for Catastrophic Loss Reduction*.
- Alsaqqar, A., B. Khudair, and R.K. Jbbar. 2017. "Rigid Trunk Sewer Deterioration Prediction Models Using Multiple Discriminant and Neural Network Models in Baghdad City, Iraq." *Journal of Engineering* 23 (8): 70–83.
- Brown, L.G. 1992. "A Survey of Image Registration Techniques." *ACM Computing Surveys* 24 (4): 535–54. [https://doi.org/10.1007/0-306-48608-3\\_13](https://doi.org/10.1007/0-306-48608-3_13).
- Buatik, A., I. Pasityothin, and K. Chaiyasarn. 2019. "Image-Based Multiview Change Detection in Concrete Structures." In *International Conference on Smart Infrastructure and Construction*, ISBN 978-0:693–99. <https://doi.org/10.1680/icsic.64669.693>.
- Calonder, M., V. Lepetit, C. Strecha, and P. Fua. 2010. "BRIEF: Binary Robust Independent Elementary Features." *Lecture Notes in Computer Science*. Vol. 6314 LNCS.  
[https://doi.org/10.1007/978-3-642-15561-1\\_56](https://doi.org/10.1007/978-3-642-15561-1_56).
- Charniak, E., and D.McDermott. 1985. "Introduction to Artificial Intiligence." published by: Addison-Wesley.de.
- Chatterjee, S., and J. S. Simonoff. 2013. *Handbook of Regression Analysis. A JOHN WILEY & SONS, INC., PUBLICATION*. <https://doi.org/10.1002/9781118532843>.

- Chen, L.C., Y. C. Shao, H. H. Jan, C. W. Huang, and Y. M. Tien. 2006. “Measuring System for Cracks in Concrete Using Multitemporal Images.” *Journal of Surveying Engineering* 132 (2): 77–82. [https://doi.org/10.1061/\(ASCE\)0733-9453\(2006\)132:2\(77\)](https://doi.org/10.1061/(ASCE)0733-9453(2006)132:2(77)).
- Choudhary, B., N. K. Sinha, and P. Shanker. 2012. “Pyramid Method in Image Processing.” *Journal of Information Systems and Communication* 3 (1): 269–73.
- Chughtai, F., and T. Zayed. 2008. “Infrastructure Condition Prediction Models for Sustainable Sewer Pipelines.” *Journal of Performance of Constructed Facilities* 22 (5): 333–41. [https://doi.org/10.1061/\(ASCE\)0887-3828\(2008\)22:5\(333\)](https://doi.org/10.1061/(ASCE)0887-3828(2008)22:5(333)).
- Clark, A., and F. Lundh. 2020. *Pillow (PIL Fork). Pillow 6.1.0 Documentation*. Vol. 8.0.0. Pillow 6.1.0 documentation. <https://pillow.readthedocs.io/en/stable/about.html>.
- Walt, S.V.D., J.L. Schönberger, J. Nunez.I., F. Boulogne, J.D. Warner, Neil Yager, E. Gouillart, and T. Yu. 2014. “Scikit-Image: Image Processing in Python.” PeerJ. Vol. Preprint. <https://doi.org/10.7717/peerj.453>.
- Daher, S. 2015. “Defect-Based Condition Assessment Model and Protocol of Sewer Pipelines.” Concordia University, Msc Thesis.
- Davari, A.A., T. Lindenberger, A. Häberle, V. Christlein, A. Maier, and C. Riess. 2017. “Image Registration for the Alignment of Digitized Historical Documents,” 1–39.
- Davies, J. P., B. A. Clarke, J. T. Whiter, and R. J. Cunningham. 2001. “Factors Influencing the Structural Deterioration and Collapse of Rigid Sewer Pipes.” *Urban Water* 3: 73–89. [https://doi.org/10.1016/S1462-0758\(01\)00017-6](https://doi.org/10.1016/S1462-0758(01)00017-6).
- Deshmukh, M., and B. Bhosle. 2011. “A Survey OF Image Registration.” *International Journal of Image Processing (IJIP)* 5 (3): 245–69.

- Duran, O., K. Althoefer, and L.D. Seneviratne. 2002. "State of the Art in Sensor Technologies for Sewer Inspection." *IEEE Sensors Journal* 2 (2): 73–81.  
<https://doi.org/10.1109/JSEN.2002.1000245>.
- El-Assaly, A., S. T. Ariaratnam, J. Ruwanpura, and H. Ng. 2006. "Cost Forecast Model for Sewer Infrastructure." *Proceedings of the Institution of Civil Engineers: Municipal Engineer* 159 (3): 155–60. <https://doi.org/10.1680/muen.2006.159.3.155>.
- Foskey, M., B. Davis, L. Goyal, S. Chang, E. Chaney, N. Strehl, S. Tomei, J. Rosenman, and S. Joshi. 2005. "Large Deformation Three-Dimensional Image Registration in Image-Guided Radiation Therapy." *Physics in Medicine and Biology* 50 (24): 5869–92.  
<https://doi.org/10.1088/0031-9155/50/24/008>.
- Gedam, A., S. Mangulkar, and B. Gandhi. 2016. "Prediction of Sewer Pipe Main Condition Using the Linear Regression Approach." *Journal of Geoscience and Environment Protection* 04: 100–105. <https://doi.org/10.4236/gep.2016.45010>.
- Gionis, Aristides, Piotr Indyk, and Rajeev Motwani. 1999. "Similarity Search in High Dimensions via Hashing." *Proceedings of the 25th International Conference on Very Large Data Bases*, 518–29. <http://dl.acm.org/citation.cfm?id=645925.671516>.
- Gobeyn, S., A.M. Mouton, A.F. Cord, A. Kaim, M. Volk, and P.L.M. Goethals. 2019. "Evolutionary Algorithms for Species Distribution Modelling: A Review in the Context of Machine Learning." *Ecological Modelling* 392: 179–95.  
<https://doi.org/10.1016/j.ecolmodel.2018.11.013>.
- Harris, C., and M. Stephens. 1988. "A Combined Corner and Edge Detector." In *Alvey Vision Conf. Oxen, UK: Alvey Vision Club*, 15:147–51. <https://doi.org/10.5244/C.2.23>.

- Hawari, A., F. Alkadour, M. Elmasry, and T. Zayed. 2020. "A State of the Art Review on Condition Assessment Models Developed for Sewer Pipelines." *Engineering Applications of Artificial Intelligence* 93 (April): 103721.  
<https://doi.org/10.1016/j.engappai.2020.103721>.
- Hyung, J., B.H. Shik, and D. Abraham. 2005. "An Ordered Probit Approach for Developing Markov Chain Based Deterioration Model for Wastewater Infrastructure Systems." In *Pipeline Division Specialty-*, 649–61. [https://doi.org/10.1061/40800\(180\)52](https://doi.org/10.1061/40800(180)52).
- Iseley, T., D. M. Abraham, and S. Gokhale. 1997. "Condition Assessment of Sewer Systems." In *Trenchless Pipeline Projects- Practical Applications*, 43–51.  
<https://doi.org/10.1061/9780784402443.006>.
- Jiang, G., J. Keller, P.L. Bond, and Z. Yuan. 2016. "Predicting Concrete Corrosion of Sewers Using Artificial Neural Network." *Water Research* 92: 52–60.  
<https://doi.org/10.1016/j.watres.2016.01.029>.
- Kabir, G., N.B.C. Balek, and S. Tesfamariam. 2018. "Sewer Structural Condition Prediction Integrating Bayesian Model Averaging with Logistic Regression." *Journal of Performance of Constructed Facilities* 32 (3): 1–10. [https://doi.org/10.1061/\(ASCE\)CF.1943-5509.0001162](https://doi.org/10.1061/(ASCE)CF.1943-5509.0001162).
- Kanan, C., and G.W. Cottrell. 2012. "Color-to-Grayscale: Does the Method Matter in Image Recognition?" *PLoS ONE* 7 (1). <https://doi.org/10.1371/journal.pone.0029740>.
- Khan, Z., T. Zayed, and O. Moselhi. 2010. "Structural Condition Assessment of Sewer Pipelines." *Journal of Performance of Constructed Facilities* 24 (2): 170–79.  
[https://doi.org/10.1061/\(ASCE\)CF.1943-5509.0000081](https://doi.org/10.1061/(ASCE)CF.1943-5509.0000081).

- Khazraeializadeh, S. 2012. "A Comparative Analysis on Sewer Structural Condition Grading Systems Using Four Sewer Condition Assessment Protocols."
- Kim, J.w., S. Kim, J. Park, and J. Nam. 2015. "Development of Crack Detection System with Unmanned Aerial Vehicles and Digital Image Processing." In *Advances in Structure Engineering and Mechanics (ASEM15)*.
- Koo, Dae., and S.T. Ariaratnam. 2006. "Innovative Method for Assessment of Underground Sewer Pipe Condition." *Automation in Construction* 15: 479–88.  
<https://doi.org/10.1016/j.autcon.2005.06.007>.
- Kuppala, K., S. Banda, and T.R. Barige. 2020. "An Overview of Deep Learning Methods for Image Registration with Focus on Feature-Based Approaches." *International Journal of Image and Data Fusion* 11 (3): 1–23. <https://doi.org/10.1080/19479832.2019.1707720>.
- Laakso, T., T. Kokkonen, I. Mellin, and R. Vahala. 2018. "Sewer Condition Prediction and Analysis of Explanatory Factors." *Water (Switzerland)* 10 (9): 1–17.  
<https://doi.org/10.3390/w10091239>.
- Li, X., F. Khademi, Y. Liu, M. Akbari, C. Wang, P.L. Bond, J. Keller, and G. Jiang. 2019. "Evaluation of Data-Driven Models for Predicting the Service Life of Concrete Sewer Pipes Subjected to Corrosion." *Journal of Environmental Management* 234 (December 2018): 431–39. <https://doi.org/10.1016/j.jenvman.2018.12.098>.
- Lin, Y., P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie. 2017. "Feature Pyramid Networks for Object Detection." In *Computer Vision and Pattern Recognition (CVPR)*, 2117–25. <https://doi.org/10.1109/ICVRIS.2019.00110>.

- Luo, C., W. Yang, P. Huang, and J. Zhou. 2019. "Overview of Image Matching Based on ORB Algorithm." *Journal of Physics: Conference Series* 1237 (3). <https://doi.org/10.1088/1742-6596/1237/3/032020>.
- Mashford, J., D. Marlow, D. Tran, and R. May. 2011. "Prediction of Sewer Condition Grade Using Support Vector Machines." *Journal of Computing in Civil Engineering* 25 (4): 283–90. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000089](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000089).
- Meeker, E. 1971. "The Improving Health of the United States, 1850-1915." *Explorations in Economic History* 9 (4): 353–73. [https://doi.org/10.1016/0014-4983\(71\)90066-0](https://doi.org/10.1016/0014-4983(71)90066-0).
- Mellett, J.S. 1995. "Ground Penetrating Radar Applications in Engineering, Environmental Management, and Geology." *Journal of Applied Geophysics* 33 (1–3): 157–66. [https://doi.org/10.1016/0926-9851\(95\)90038-1](https://doi.org/10.1016/0926-9851(95)90038-1).
- Mirza, S. 2007. "Danger Ahead: The Coming Collapse Of Canada's Municipal Infrastructure". Report to: Federation of Canadian Municipalities.
- Mistry, K., and A. Saluja. 2016. "An Introduction to OpenCV Using Python with Ubuntu." *International Journal of Scientific Research in Computer Science, Engineering and Information Technology* 1 (2): 65–68. <http://opencv.org>.
- Mohammadi, M. M. 2019. "Development of Condition Models for Sanitary Sewer Pipes." *University of Texas at Arlington*. University Of Texas At TEXAS AT Arlington, Ph.D.
- Najafi, M., and G. Kulandaivel. 2005. "Pipeline Condition Prediction Using Neural Network Models." In *Pipeline Division Specialty Confedence*, 767–81. [https://doi.org/10.1061/40800\(180\)61](https://doi.org/10.1061/40800(180)61).

- Navab-kashani, R. 2014. "Productivity Analysis of Closed Circuit Television ( CCTV ) Sewer Mainline Inspection." University of Alberta, MSc Thesis.
- Oliveira, F.P.M., and J.M.R.S. Tavares. 2014. "Medical Image Registration: A Review." *Computer Methods in Biomechanics and Biomedical Engineering* 17 (2): 73–93. <https://doi.org/10.1201/b15511>.
- Padmavathi, K., and K. Thangadurai. 2016. "Implementation of RGB and Grayscale Images in Plant Leaves Disease Detection - Comparative Study." *Indian Journal of Science and Technology* 9 (6): 4–9. <https://doi.org/10.17485/ijst/2016/v9i6/77739>.
- Park, T. 2009. "A Comprehensive Asset Management System For Sewer Infrastructures." *The Pennsylvania State University, Ph.D. Thesis*. <https://doi.org/10.1029/2008GB003237>.
- Rosin, P.L. 1999. "Measuring Corner Properties." *Computer Vision and Image Understanding* 73 (2): 291–307. <https://doi.org/10.1006/cviu.1998.0719>.
- Rosten, E., and T. Drummond. 2005. "Fusing Points and Lines for High Performance Tracking." In *IEEE International Conference on Computer Vision*, II:1508–15. <https://doi.org/10.1109/ICCV.2005.104>.
- Rosten, E., and T. Drummond. 2006. "Machine Learning for High-Speed Corner Detection." *Lecture Notes in Computer Science*. Vol. 3951. [https://doi.org/10.1007/11744023\\_34](https://doi.org/10.1007/11744023_34).
- Russell, S. J., and P. Norvig. 2010. "Artificial intelligence: A modern approach." Boston: Pearson
- Ruble, E., V. Rabaud, K. Konolige, and G. Bradski. 2011. "ORB: An Efficient Alternative to SIFT or SURF." In *2011 International Conference on Computer Vision*, 2564–71. IEEE.

<https://doi.org/10.1109/ICCV.2011.6126544>.

Saravanan, C. 2010. "Color Image to Grayscale Image Conversion." In *International Conference on Computer Engineering and Applications, ICCEA*, 2:196–99.

<https://doi.org/10.1109/ICCEA.2010.192>.

Sathya, R., and A. Abraham. 2013. "Comparison of Supervised and Unsupervised Learning Algorithms for Pattern Classification." *International Journal of Advanced Research in Artificial Intelligence* 2 (2): 34–38. <https://doi.org/10.14569/ijarai.2013.020206>.

Scheidegger, A., T. Hug, J. Rieckermann, and M. Maurer. 2011. "Network Condition Simulator for Benchmarking Sewer Deterioration Models." *Water Research* 45 (16): 4983–94.

<https://doi.org/10.1016/j.watres.2011.07.008>.

Sibbald, A., Fairfield, C.A. and Forde, M.C. (1995). Impulse Testing of Brickwork Sewers, *Journal of Low Frequency Noise & Vibration*, 1995, Vol. 14, No. 1, pp 43-54.

Siddharth, S., and K. Rajeev. 2014. "A Survey of Recent and Classical Image Registration Methods." *International Journal of Signal Processing, Image Processing and Pattern Recognition* 7 (4): 167–76. <https://doi.org/2005-4254> IJSIP

Sohn, H.G., Y.M. Lim, K.H. Yun, and G.H. Kim. 2005. "Monitoring Crack Changes in Concrete Structures." *Computer-Aided Civil and Infrastructure Engineering* 20 (1): 52–61.

<https://doi.org/10.1111/j.1467-8667.2005.00376.x>.

Song, Z., S. Zhou, and J. Guan. 2014. "A Novel Image Registration Algorithm for Remote Sensing under Affine Transformation." *IEEE Transactions on Geoscience and Remote Sensing* 52 (8): 4895–4912. <https://doi.org/10.1109/TGRS.2013.2285814>.

- Sousa, V., J.P. Matos, and N. Matias. 2014. "Evaluation of Artificial Intelligence Tool Performance and Uncertainty for Predicting Sewer Structural Condition." *Automation in Construction* 44: 84–91. <https://doi.org/10.1016/j.autcon.2014.04.004>.
- Syachrani, S., H.S. Jeong, and C.S. Chung. 2013. "Decision Tree-Based Deterioration Model for Buried Wastewater Pipelines." *Journal of Performance of Constructed Facilities* 27 (5): 633–45. [https://doi.org/10.1061/\(ASCE\)CF.1943-5509.0000349](https://doi.org/10.1061/(ASCE)CF.1943-5509.0000349).
- Tareen, S.A.K., and Z. Saleem. 2018. "A Comparative Analysis of SIFT, SURF, KAZE, AKAZE, ORB, and BRISK." In *International Conference on Computing, Mathematics and Engineering Technologies — ICoMET*, 1–10. IEEE. <https://doi.org/10.1109/ICOMET.2018.8346440>.
- Toffol, S.D., C. Engelhard, and W. Rauch. 2007. "Combined Sewer System versus Separate System - A Comparison of Ecological and Economical Performance Indicators." *Water Science and Technology* 55 (4): 255–64. <https://doi.org/10.2166/wst.2007.116>.
- Tran, D. H., A. W.M. Ng, B. J.C. Perera, S. Burn, and P. Davis. 2006. "Application of Probabilistic Neural Networks in Modelling Structural Deterioration of Stormwater Pipes." *Urban Water Journal* 3 (3): 175–84. <https://doi.org/10.1080/15730620600961684>.
- Tran, D.H. 2007. "Investigation of Deterioration Models for Stormwater Pipe Systems." *Victoria University, PhD Thesis*.
- Wirahadikusumah, R., D. Abraham, and T. Iseley. 2001. "Challenging Issues In Modeling Deterioration Of Combined Sewers." *JOURNAL OF INFRASTRUCTURE SYSTEMS* 7 (2): 77–48.

- Wyawahare, M.V., P.M. Patil, and H.K. Abhyankar. 2009. "Image Registration Techniques : An Overview." *International Journal of Signal Processing, Image Processing and Pattern Recognition* 2 (3): 11–28.
- Yang, Y.S., C.I. Wu, T.T.C. Hsu, H.C. Yang, H.J. Lu, and C.C. Chang. 2018. "Image Analysis Method for Crack Distribution and Width Estimation for Reinforced Concrete Structures." *Automation in Construction* 91: 120–32. <https://doi.org/10.1016/j.autcon.2018.03.012>.
- Yin, X., Y. Chen, A. Bouferguene, H. Zaman, M. Al-Hussein, and L. Kurach. 2020. "A Deep Learning-Based Framework for an Automated Defect Detection System for Sewer Pipes." *Automation in Construction* 109 (June). <https://doi.org/10.1016/j.autcon.2019.102967>.
- Yin, X., Y. Chen, A. Bouferguene, H. Zaman, M. Al-Hussein, R. Russell, and L. Kurach. 2019. "Standard Closed-Circuit Television (CCTV) Collection Time Extraction of Sewer Pipes with Machine Learning Algorithm." In *Proceedings of the 36th International Symposium on Automation and Robotics in Construction, ISARC.*, 107–13. <https://doi.org/10.22260/isarc2019/0015>.
- Zelinski, M.E., J.R. Henderson, and E. Held. 2018. "Image Registration and Change Detection for Artifact Detection in Remote Sensing Imagery." In *The Society of Photo-Optical Instrumentation Engineers*. <https://doi.org/10.1117/12.2303934>.
- Zhao, J.Q. 1998. "APWA International Public Works Congress NRCC/CPWA Seminar Series &quot;Innovations in Urban Infrastructure&quot;; 75 TRUNK SEWERS IN CANADA." In *1998 APWA International Public Works Congress Seminar Series, American Public Works Association. Las Vegas*, 75–89. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.3.1681&rep=rep1&type=pdf>.

Zitová, B., and J. Flusser. 2003. "Image Registration Methods: A Survey." *Image and Vision Computing* 21 (11): 977–1000. [https://doi.org/10.1016/S0262-8856\(03\)00137-9](https://doi.org/10.1016/S0262-8856(03)00137-9).