

# Computer Vision Techniques for Automatic Structural Assessment of Underground Pipes

Sunil K. Sinha\*

Department of Civil and Environmental Engineering, Pennsylvania State University, University Park, PA 16803, USA

Paul W. Fieguth

Department of Systems Design Engineering, University of Waterloo, Ontario, Canada N2L 3G1

&

Maria A. Polak

Department of Civil Engineering, University of Waterloo, Ontario, Canada N2L 3G1

**Abstract:** Pipeline surface defects such as cracks cause major problems for asset managers, particularly when the pipe is buried under the ground. The manual inspection of surface defects in the underground pipes has a number of drawbacks, including subjectivity, varying standards, and high costs. An automatic inspection system using image processing and artificial intelligence techniques can overcome many of these disadvantages and offer asset managers an opportunity to significantly improve quality and reduce costs. This article presents a system for the application of computer vision techniques to the automatic assessment of the structural condition of underground pipes. The algorithm consists of image preprocessing, a sequence of morphological operations to accurately extract pipe joints and laterals (where smaller pipe is connected to main bigger pipe), and statistical filters for detection of surface cracks in the pipeline network. The proposed approach can be completely automated and has been tested on over 1,000 scanned images of underground pipes from major cities in North America.

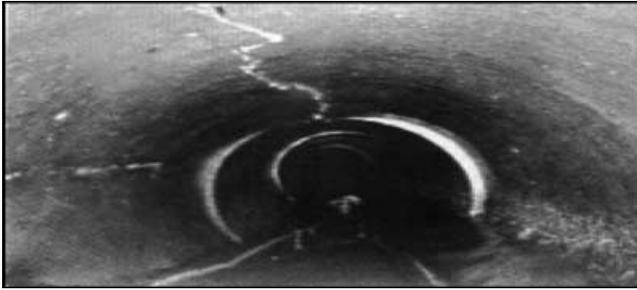
\*To whom correspondence should be addressed. E-mail: sunil@engr.psu.edu.

## 1 INTRODUCTION

Beneath North America's roads lie thousands of miles of pipe that bring purified water to homes and carry away wastewater (sewage and storm water). For the most part, these systems have been functioning longer than their intended design life (i.e., 50 years for concrete pipe), with little or no repair. Hence, they are in a state of deterioration.

Maintenance and rehabilitation (M&R) of pipeline systems pose a major challenge for most municipalities in North America, given their budgetary constraints, the demand on providing quality service, and the need for preserving their pipeline infrastructure. Neglecting regular M&R of these underground pipelines adds to life-cycle costs and liabilities and, in extreme cases, causes stoppage or reduction of vital services.

Accurate pipeline condition assessment is vital to developing a cost-effective and efficient pipeline M&R program. At present, the assessed condition of underground pipes is based on the subjective visual inspection of closed circuit television (CCTV) surveys (Iseley et al., 1997). CCTV surveys are conducted using a remotely controlled vehicle carrying a television camera



**Fig. 1.** Typical images of underground pipe scanned by closed circuit television (CCTV) camera in the city of Toronto.

through an underground pipe. The data acquired from this process consist of videotape, photographs of specific defects, and a record produced by the technician. A typical scanned image of CCTV surveys is shown in Figure 1. Diagnosis of defects depends on the experience, capability, and concentration of the operator, making the detection of defect error prone. A large number of new technologies such as pipe scanner and evaluation technology (PSET) (Iseley, 1999), laser-based systems (Van Cauwelaert et al., 1989), etc. have made it possible to obtain high-quality images of pipes. PSET is an innovative technology for obtaining unfolded images of the interior of pipes (Iseley, 1999). This is accomplished by utilizing scanner and gyroscope technology. A typical scanned image of PSET surveys is shown in Figure 2. Although underground imaging technology has made substantial strides in recent years, the basic means of analysis are unchanged: a technician is required to identify defects on a television monitor. The research of this article seeks to address this latter limitation, thus allowing the technician to do exactly what he has been trained to do, which is to insure that the inspection equipment is being operated properly.

This article addresses the development of an automated underground pipe inspection system. Emphasis has been placed on investigations of algorithms and techniques for image processing, feature extraction, and



**Fig. 2.** Typical images of underground pipe scanned by pipe scanner and evaluation technology (PSET) camera in the city of Toronto.

pattern classification. In particular, this research has explored how various signals and image processing concepts, nonlinear filtering, feature extraction, pattern classification, and artificial intelligence techniques can be judiciously synthesized for computationally efficient and robust identification of underground pipe defects (i.e., cracks, holes, collapse surface, and defective joints and laterals). The proposed automated system could overcome many of the limitations of the current manual inspection of pipes and can provide a more accurate assessment of underground pipe conditions.

## 2 METHODS OF INSPECTING INNER SURFACE OF PIPES

Nondestructive testing (NDT) (Davies and Mamlouk, 1985) is the branch of engineering concerned with non-contact methods of detecting and evaluating defects in materials. Defects can affect the serviceability of the material or structure, so NDT is important in guaranteeing safe operation as well as in quality control and assessing pipe life. The defect may be cracks or inclusions in welds and castings, or variations in structural properties that can lead to loss of strength or failure in service. Nondestructive testing is used for in-service inspection and for condition monitoring. It is also used for measurement of components and for the measurement of physical properties such as hardness and internal stress. The essential feature of NDT is that the test process itself produces no deleterious effects on the material or structure under test. The subject of NDT has no clearly defined boundaries; it ranges from simple techniques such as visual examination of surfaces, through the well-established methods of radiography, ultrasonic testing, magnetic particle crack detection, to new and very specialized methods. NDT methods can be adapted to automate production processes as well as to the inspection of localized problem areas. All NDT techniques have the ability to measure only specific types of defects, material properties, and/or material response. Therefore, the best choice of an NDT method in a specific pipeline application will depend on the pipeline physical properties and defects. Thus, before the selection of an appropriate NDT method, a thorough knowledge of each NDT method application and its limitations is required along with a good understanding of the piping system.

### 2.1 Non-visual NDT methods

**2.1.1 Ultrasonic inspection (sonar).** Ultrasonic inspection (Birks and Green, 1991) is performed using a beam of very high frequency coherent sound energy, with the frequency being many orders of magnitude higher than

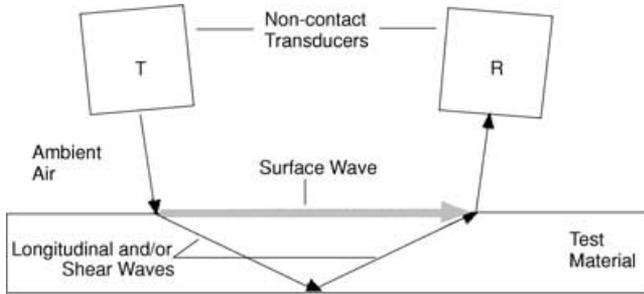


Fig. 3. Ultrasonic inspection of pipes.

a human being can hear. The sound wave travels into the object being inspected and reflects whenever there is a change in the density of the material, with some of the energy in the wave returning to the surface and some passing on through the new material. The ultrasonic inspection process is shown in Figure 3. Ultrasonic beams can be used to image the human body, inspect aircraft, or examine oil pipelines. The technique is capable of detecting pits, voids, and cracks, although certain crack orientations are much more difficult to detect than others. The ultrasonic wave reflects most easily when it crosses an interface between two materials that are perpendicular to the wave.

**2.1.2 Eddy current testing.** Eddy current testing (Joynson et al., 1986) is an electromagnetic technique that can detect surface and subsurface discontinuities in tube walls up to about 3/8" (10 mm) thick on conductive materials. Applications range from crack detection, to the rapid sorting of small components for defects, size variations, or material variation. When an energized coil is brought near the surface of a metal component, eddy currents are induced into the specimen. These currents set up a magnetic field that tends to oppose the original magnetic field. The impedance of the coil in close proximity to the specimen is influenced by the presence of the induced eddy currents in the specimen. When the eddy currents in the specimen are distorted by the presence of a defect or a material variation, the impedance in the coil is altered. This change is measured and displayed in a manner that indicates the type of defect or material condition. This method is commonly performed on heat exchanger tubing by inserting a probe down the full length of each tube to be inspected as shown in Figure 4. The probe contains a coil arrangement, energized by alternating currents operating at one or more frequencies. The electrical impedance of the test coil arrangement is modified by the proximity of the tube, tube dimensions, electrical conductivity, magnetic permeability of the tube material, and metallurgical and mechanical discontinuities. Wear on the tube surface

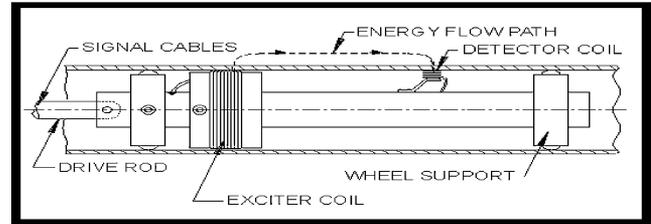


Fig. 4. Eddy current inspection of pipes.

under a support is also detectable. The electromagnetic response caused by passing these variables produces electrical signals, which are processed electronically to produce a visual response characteristic of the change encountered. Visual responses, often called "signatures," are displayed on the test instrument monitor for evaluation by the field technician (analyst).

**2.1.3 Acoustic emission monitoring.** This method involves listening to the sounds (which are usually inaudible to the human ear) made by a material, structure, or machine in use or under load (Bassim and Houssny-Emam, 1983). Conclusions are made about its "state of health" from what is heard, just like the doctor who listens to your heart and lungs. The technique involves attaching one or more ultrasonic microphones to the object and analyzing the sounds using computer-based instruments. The noises may arise from friction (including bearing wear), crack growth, turbulence (including leakage), and material changes such as corrosion. Applications include testing pipelines and storage tanks (above and below the ground), fiberglass structures, rotating machinery, weld monitoring, and biological and chemical changes.

## 2.2 Visual NDT methods

Visual inspection is an NDT method used extensively to evaluate the condition or the quality of a component (Krstulovic et al., 1996). It is easy to perform, inexpensive, and usually does not require special equipment. It is most effective for the inspection of welds where quick detection and the correction of defects or process-related problems can result in significant cost savings. It is the primary evaluation method of many quality control programs. The method requires good vision, good lighting, and operator knowledge.

Most municipal pipeline systems are inspected visually by mobile closed circuit television (CCTV) systems or human inspectors. CCTV examination using a mobile camera system is the typical approach to this type of examination. However, there are several CCTV variants that may reduce the cost of the inspection or provide

improved results. There are also alternative techniques that will work where CCTV will not or that will give direct measurements of pipe condition as opposed to the estimates produced by CCTV inspection. Each inspection technique is discussed separately below.

**2.2.1 CCTV inspection.** There are two basic types of CCTV system (Wirahadikusumah et al., 1998). Each uses a television camera in conjunction with a video monitor, videocassette recorders, and possibly other recording devices. In one case the inspection is performed using a stationary or zoom camera mounted at a manhole so that it looks into the pipe, whereas in the other a mobile, robotic system is placed within the pipe itself. The camera provides images to an operator who is trained to detect, classify, and rate the severity of defects against documented criteria. The typical CCTV camera scanning process of underground pipes is shown in Figure 5. Either form of CCTV inspection may miss certain types of defects, especially those that are hidden from the camera by obstructions as it looks down the pipe. This method is also vulnerable to lapses in operator concentration, inexperience, and the inability of the image to reveal important defects (Wirahadikusumah et al., 1998). Thus, the results are widely agreed to lack consistency and the reliability to track deterioration so that preventive maintenance can be undertaken with confidence. It does, however, provide useful information on gross defects.



**Fig. 5.** CCTV camera inspection process of underground concrete pipe.

**2.2.2 Pipe scanner and evaluation technology.** Pipe scanner and evaluation technology (PSET) is an innovative technology for obtaining images of the interior of pipe. PSET was developed by TOA Grout, CORE Corp., California, and the Tokyo Metropolitan Government's Services (TGS) Company. PSET is a system that offers a new inspection method minimizing some of the shortcomings of the traditional inspection equipment that relies on a CCTV inspection. This is accomplished by utilizing scanning and gyroscopic technology. The mechanics of inspecting the pipes by PSET camera are similar to the CCTV inspection. The PSET is designed to operate from a tractor platform to propel the tool through the pipe. Since the PSET utilizes state-of-the-art scanner technology, it can travel through the pipe at a uniform rate of speed. The major benefit of the PSET system over the current CCTV technology is that the engineer will have higher quality image data to make critical rehabilitation decisions.

**2.2.3 Laser-based scanning systems.** In addition to the simple light line system described above, lasers have been used in the past to evaluate both the shapes of pipelines and the types of defects they contain (Hibino et al., 1994). These systems are restricted to the part of the pipe above the waterline, but they can, in theory, make possible extremely accurate inspections of pipe condition. An additional advantage to this approach is that the information from the laser scans is readily recorded and analyzed by computer, substantially reducing operator errors. Although the initial equipment may be more expensive than a CCTV system, the reduced operator time necessary to use the technique may also mean that its operation will be more economical. The technology is still in the development stage.

### 3 METHODOLOGY FOR AUTOMATIC IMAGE-BASED INSPECTION

Industry is increasingly using machine vision systems to aid in the manufacturing and quality-control processes (Newman and Jain, 1995). The goal of a machine vision (Chin and Harlow, 1982) is to create a model of the real world from images. A machine vision system recovers useful information about a scene from its two-dimensional projections. Since images are two-dimensional projections of the three-dimensional world, the information is not directly available and must be recovered. To recover the information, knowledge about the objects in the scene is required. The emphasis in machine vision systems is on maximizing automatic operation at each stage, and these systems should use knowledge to accomplish this.

Machine vision emulates human vision in that it attempts to interpret images. Human vision deals with the global information available in a scene, resolves ambiguities due to perspective, lighting, and attribute, and can perform guidance through unfamiliar territories. In fact, human vision has been shown to be incapable of performing reliable inspection (Agin, 1980), since the human vision process is prone to subjective considerations, fatigue, and boredom, which interfere with consistent evaluations. Also, the human vision is limited to the visible spectrum, while a machine vision system can exploit a much larger range of the electromagnetic spectrum, including infrared radiation, X-rays, and ultrasounds, thereby making it suitable for a wide range of nondestructive testing and inspection process-related tasks (Ballard and Brown, 1982).

An automated image-based inspection aims to extract information from an image on the conditions of objects represented in the image. Usually it is impossible to extract this information concerning the dimensions of objects or their defect properties directly from the image. Basically, all machine vision systems involve image acquisition, image preprocessing, segmentation, and extracting relevant features for classification of the type, severity, and extent of defects present in the image.

### 3.1 Image processing and segmentation

Vision allows humans to perceive and understand the world surrounding them. Computer vision aims to duplicate the effect of human vision by electronically perceiving and understanding an image. Giving computers the ability to see is not an easy task. We live in a three-dimensional (3D) world, and when computers try to analyze objects in 3D space, the visual sensors available (e.g., TV cameras) usually give two-dimensional (2D) images, and this projection to a lower number of dimensions incurs an enormous loss of information.

Image preprocessing and segmentation is the initial stage for any recognition process, whereby the acquired image is “broken up” into meaningful regions or segments. The segmentation process is not primarily concerned with what the regions represent, but only with the process of partitioning the image. In the simplest case (binary images) there are only two regions: a foreground (object) region and a background region. In gray level images, there may be many types of regions or classes within the image; for example, in a natural scene to be segmented, there may be regions of sky, clouds, ground, building, and trees. There are, broadly speaking, two approaches to image segmentation, namely, thresholding and region- or edge-based methods (Pratt, 1978).

**3.1.1 Image preprocessing.** The principal objective of image preprocessing is to process an image so that the result is more suitable than the original image for a specific application. The word *specific* is important, because it establishes at the outset that the techniques discussed in this section are very much problem-oriented. Thus, for example, a method that is quite useful for enhancing X-ray images may not necessarily be the best approach for enhancing images of underground pipes.

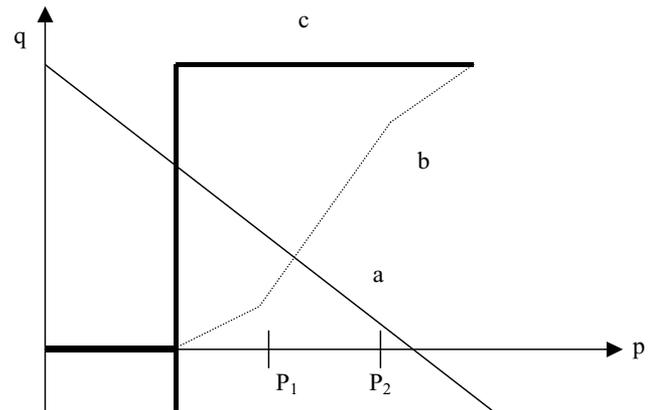
#### 3.1.1.1 Gray scale transformation

Gray scale transformations (Rosenfeld and Kak, 1982) do not depend on the position of the pixel in the image. A transformation  $\lambda$  of the original brightness  $p$  from scale  $[q_o, q_k]$  into brightness  $q$  from a new scale  $[p_o, p_k]$  is given by

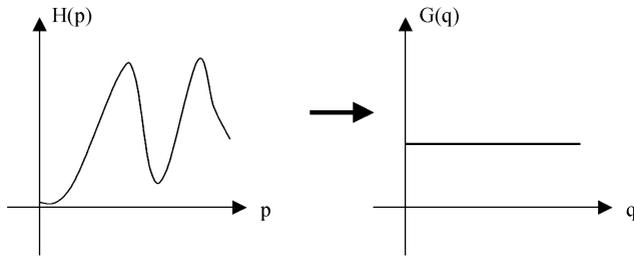
$$q = \lambda(p) \quad (1)$$

The most common gray scale transformations are shown in Figure 6; the straight line  $a$  denotes the negative transformation; the piecewise linear function  $b$  enhances the image contrast between brightness values  $p_1$  and  $p_2$ . The function  $c$  is called brightness thresholding and results in a black-and-white image.

A gray scale transformation for contrast enhancement is usually found automatically using the histogram equalization technique (Rosenfeld and Kak, 1982). The aim is to create an image with equally distributed brightness levels over the whole brightness scale (see Figure 7). Histogram equalization enhances contrast for brightness values close to histogram maxima and decreases contrast near minima.



**Fig. 6.** Some gray-scale transformations. The straight line  $a$  denotes the negative transformation; the piecewise linear function  $b$  enhances the image contrast between brightness values  $p_1$  and  $p_2$ . The function  $c$  is called *brightness thresholding*.



**Fig. 7.** Histogram equalization. The aim is to create an image with equally distributed brightness levels over the whole brightness scale.

### 3.1.1.2 Image smoothing

Image smoothing (Castleman, 1996) is the set of local processing methods whose predominant use is the suppression of image noise. Calculation of the new value is based on the averaging of brightness values in some neighborhood. Smoothing poses the problem of blurring sharp edges in the image, and so we shall concentrate on smoothing methods that are edge preserving. They are based on the general idea that the average is computed only from those points in the neighborhood that have similar properties to the point being processed.

### 3.1.1.3 Color image processing

The use of color in image processing (Smith, 1978) is motivated by two principal factors. First, in automated image analysis, color is a powerful descriptor that often simplifies object identification and extraction from a scene. Second, in image analysis performed by human beings, the motivation for color is that the human eye can discern thousands of color shades and intensities, compared to about only two dozen shades of gray (Smith, 1978). Color is connected with the ability of objects to reflect electromagnetic waves of different wavelengths; the chromatic spectrum spans the electromagnetic spectrum from approximately 400 nm to 700 nm. Humans detect colors as combinations of the primary colors red, green, and blue, which for the purpose of standardization have been defined as 700 nm, 546.1 nm, and 435.8 nm, respectively (Smith, 1978), although this standardization does not imply that all colors can be synthesized as combinations of these three.

The purpose of a color model is to facilitate the specification of colors in some standard, generally accepted way. In essence, a color model is a specification of a 3D coordinate system and a subspace within that system where each color is represented by a single point. The color models most often used are the RGB, the YIQ, and the HIS models (Gonzalez and Wintz, 1987).

**3.1.2 Image segmentation.** Image segmentation is one of the most important steps leading to the analysis of processed image data—its main goal is to divide an image into parts that have a strong correlation with objects or areas of the real world contained in the image. Image data ambiguity is one of the main segmentation problems, often accompanied by information noise. Segmentation methods can be divided into three groups according to the dominant features they employ. First is global knowledge about an image or its part; the knowledge is usually represented by a histogram of image features (Weszka and Rosenfeld, 1979). Edge-based segmentations form the second group (Kundu and Mitra, 1987), and region-based segmentations the third—many different characteristics may be used in edge detection or region growing (Chang and Li, 1995), for example, brightness, texture, velocity field, etc. The second and the third groups solve a dual problem. Each region can be represented by its closed boundary, and each closed boundary describes a region. Because of the different natures of the various edge- and region-based algorithms, they may be expected to give somewhat different results and consequently different information.

### 3.1.2.1 Threshold-based segmentation

Gray level thresholding (Chow and Kaneko, 1972) is the simplest segmentation process. Many objects or image regions are characterized by constant reflectivity or light absorption of their surfaces; a brightness constant or threshold can be determined to segment objects and background. Thresholding is computationally inexpensive and fast—it is the oldest segmentation method and is still widely used in simple applications (Chow and Kaneko, 1972).

A complete segmentation of an image  $R$  is a finite set of regions  $R_1, \dots, R_S$ ,

$$R = \bigcup_{i=1}^S R_i \quad R_i \cap R_j = \emptyset \quad i \neq j \quad (2)$$

Segmentation can result from thresholding in simple scenes. Thresholding is the transformation of an input image  $f$  to an output (segmented) binary image  $g$  as follows:

$$g(i, j) = 1 \quad \text{if } f(i, j) \geq T \\ = 0 \quad \text{if } f(i, j) < T \quad (3)$$

where  $T$  is the threshold,  $g(i, j) = 1$  for image elements of objects, and  $g(i, j) = 0$  for image elements of the background (or vice versa). If objects do not touch each other, and if their gray levels are clearly distinct from background gray levels, thresholding is a suitable segmentation method.

Correct threshold selection is crucial for successful threshold segmentation; this selection can be determined interactively or it can be the result of some threshold detection method. Only under very unusual circumstances can thresholding be successful using a single threshold for the whole image (global thresholding), since even in very simple images there are likely to be gray level variations in objects and background; this variation may be due to nonuniform lighting, nonuniform input device parameters, or a number of factors. Segmentation using variable thresholds (also called adaptive thresholding), in which the threshold value varies over the image as a function of local image characteristics, can produce the solution in these cases.

Methods based on approximation of the histogram of an image using a weighted sum of two or more probability densities with normal distribution represent a different approach called optimal thresholding (Chow and Kaneko, 1972). The threshold is set as the closest gray level corresponding to the minimum probability between the maxima of two or more normal distributions, which results in minimum error segmentation. The difficulty with these methods is in estimating normal distribution parameters together with the uncertainty that the distribution may be considered normal. These difficulties may be overcome if an optimal threshold is sought that maximizes gray level variance between objects and background (Kittler and Illingworth, 1985).

### 3.1.2.2 Edge-based segmentation

Edge-based segmentation represents a large group of methods based on information about edges in the image; it is one of the easiest segmentation approaches and still remains very important. Edge-based segmentations rely on edges found in an image by edge-detecting operators—these edges mark image locations of discontinuities in gray level, color, texture, etc. The image resulting from edge detection cannot be used as a segmentation result; supplementary processing steps must follow to combine edges into edge chains that correspond better with borders in the image. The most common problems of edge-based segmentation, caused by image noise or unsuitable information in an image, are an edge presence in locations where there is no border, and no edge presence where a real border exists. Clearly both these cases have a negative influence on segmentation results.

Edge detectors are a collection of very important local image segmentation methods used to locate changes in the intensity function; edges are pixels where this function (brightness) changes abruptly. Calculus describes changes of continuous functions using derivatives; an image function depends on two variables—coordinates in the image plane—and so operators describing edges are

expressed using partial derivatives. A change of the image function can be described by a gradient that points in the direction of the largest growth of the image function. An edge is a property attached to an individual pixel and is calculated from the image function behavior in a neighborhood of that pixel. It is a vector variable with two components, magnitude and direction. The edge magnitude is the magnitude of the gradient, and the edge direction  $\phi$  is rotated with respect to the gradient direction  $\psi$  by  $-90^\circ$ . The gradient direction gives the direction of maximum growth of the function, e.g., from black [ $f(i, j) = 0$ ] to white [ $f(i, j) = 255$ ].

The gradient magnitude  $|\text{grad } g(x, y)|$  and gradient direction  $\psi$  are continuous image functions calculated as

$$|\text{grad } g(x, y)| = \sqrt{\left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2} \quad (4)$$

$$\psi = \arg\left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}\right)$$

where  $\arg(x, y)$  is the angle (in radians) from the  $x$ -axis to the point  $(x, y)$ .

Edge detection represents an extremely important step facilitating higher-level image analysis and therefore remains an area of active research, with new approaches continually being developed. Recent examples include edge detectors using fuzzy logic (Law et al., 1996), neural networks (Vrabel, 1996), or wavelets (Aydin et al., 1996). It may be difficult to select the most appropriate edge detection strategy; a comparison of edge detection approaches and an assessment of their performance may be found in Ramesh and Haralick (1994) and Demigny et al. (1995).

### 3.1.2.3 Region-based segmentation

The aim of the segmentation methods described in the previous section was to find borders between regions; the methods discussed in this section construct regions directly. It is easy to construct regions from their borders, and it is easy to detect borders of existing regions. However, segmentations resulting from edge-based methods and region-growing methods are not usually exactly the same, and a combination of results may often be a good idea. Region-growing techniques are generally better in noisy images, where borders are extremely difficult to detect. Homogeneity is an important property of regions and is used as the main segmentation criterion in region growing, whose basic idea is to divide an image into zones of maximum homogeneity. The criteria for homogeneity can be based on gray level, color, texture, shape, model (using semantic information), etc. (Haralick and Shapiro, 1985; Pal and Pal, 1987). Properties chosen to describe regions influence the form,

complexity, and amount of prior information in the specific region-growing segmentation method. Methods that specifically address region-growing segmentation of color images are reported in Priese and Rehrmann (1993) and Schettini (1993).

### 3.2 Feature extraction

The previous section was devoted to image segmentation methods and showed how to construct homogeneous regions of images and/or their boundaries. Recognition of image regions is an important step on the way to understanding image data and requires an exact region description in a form suitable for a classifier. This description should generate a numeric feature vector, which characterizes properties (e.g., shape) of the region.

If a system to distinguish objects of different types is desired, then it is important to first decide which characteristics of the objects should be measured to produce descriptive parameters. The particular characteristics that are measured are called the features of the object, and the resulting parameter values comprise the feature vector for each object. Proper selection of the features is important, since only these will be used to identify the objects. There are few analytical means to guide the selection of features. Frequently, intuition guides the listing of potentially useful features. Feature-ordering techniques (Castleman, 1996) compute the relative power of the various features. In practice, the feature selection process usually involves testing a set of intuitively reasonable features and reducing the set to an acceptable number of the best ones. Good features have four characteristics (Castleman, 1996):

- *Discrimination*: Features should take on significantly different values for objects belonging to different classes.
- *Reliability*: Features should take on similar values for all objects of the same class.
- *Independence*: The various features used should be uncorrelated with each other.
- *Small number*: The complexity of a pattern recognition system increases rapidly with the dimensionality (number of features used) of the system. More importantly, the number of objects required to train the classifier and to measure its performance increases exponentially with the number of features (Jain and Chandrasekeran, 1982). In some cases, it may be impractical to acquire the amount of data required to train the classifier adequately. Finally, adding more features that are either noisy or highly correlated with existing features can actually degrade the performance of the classifier, particularly in view of the lim-

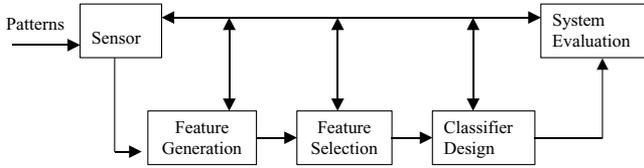
ited size of the training set (Kanal and Chandrasekeran, 1971).

In practice, the feature selection process usually involves testing a set of intuitively reasonable features and reducing the set to an acceptable number of the best ones. Many features can be used to describe an object. The most basic of all image features is the measure of image amplitude in terms of spectral value, luminance, tristimulus value, or other units (Castleman, 1996). Image transforms provide the frequency domain information in the data. Transform coefficient feature extraction has proved practical in several applications in which the transform domain features are used as inputs to a pattern recognition classification system (Haralick, 1973; Shaikh and Tian, 1996). The textural features of an object can often be used to discriminate between the surface finish of a smooth or coarsely textured object (Shaikh and Tian, 1996). The gray level histogram of an image often contains sufficient information to allow analysis of the image content and, in particular, to discriminate between objects (Shaikh and Tian, 1996). The most common object measurements made are those that describe shape.

Defining the shape of an object can prove to be very difficult. Shape is usually represented verbally or in figures, and people use terms such as *elongation*, *rounded*, *with sharp edges*, etc. The computer era has introduced the necessity to describe even very complicated shapes precisely, and while many practical shape description methods exist, there is no generally accepted methodology of shape description. Further, it is not known what makes a shape important. In general, shape descriptors are sorted according to whether they are based on object boundary information (e.g., contour-based, external description) or whether the information from object regions is used (e.g., region-based, internal description) (Hogg, 1993). This classification of shape description methods may be local or global and differ in sensitivity to translation, rotation, scaling, etc.

### 3.3 Pattern recognition

Pattern recognition is the scientific discipline whose goal is the classification of objects into a number of categories or classes. Depending on the application, these objects can be images, signal waveforms, or any type of measurements that need to be classified. Pattern recognition has a long history, but before the 1960s, it was mostly the output of theoretical research in the area of statistics. As with everything else, the advent of computers increased the demand for practical applications of pattern recognition, which in turn set new demands for further theoretical developments. The theory of pattern recognition is thoroughly discussed in several references (Duda and



**Fig. 8.** The basic stages involved in the design of a pattern classification system.

Hart, 1970; Fu, 1982; Oja, 1983; Schalkoff, 1992), but here only a brief introduction will be given.

Machine vision is an area in which pattern recognition is of importance. For example, in inspection, manufactured objects on a moving conveyor may pass the inspection station, where the camera stands, and it has to be ascertained whether there is a defect. Thus, images have to be analyzed, and a pattern recognition system has to classify the objects.

An object is a physical unit, usually represented in image analysis and computer vision by a region in a segmented image. The set of objects can be divided into disjoint subsets, which, from the classification point of view, have some common features and are called *classes*. The definition of how the objects are divided into classes is ambiguous and depends on the classification goal. The classifier (similarly to a human) does not decide about the class from the object itself; rather, sensed object properties serve this purpose. For example, to distinguish steel from sandstone, their molecular structures do not have to be determined, although this would describe these materials well. Properties such as texture, specific weight, hardness, etc., are used instead. This sensed object is called the pattern, and the classifier does not actually recognize objects but recognizes their patterns. The main pattern recognition steps are shown in Figure 8. As is apparent from the feedback arrows, these steps are not independent. On the contrary, they are interrelated and, depending on the results, one may go back to redesign earlier stages in order to improve the overall performance.

#### 4 DEVELOPMENT OF AUTOMATED PIPE INSPECTION SYSTEM

We have acquired a data set consisting of thousands of images of underground pipes from 15 major cities in North America. This data set has been used to explore basic characteristics of underground pipe images. Analyses of images have shown that there are two important characteristics that complicate the segmentation of pipe images: first, the presence of a complicated background pattern due to earlier runoff, patches of repair work, cor-

roded areas, debris, nonuniformities in illumination, and flaws in the image acquisition process; second, the three main objects of interest—cracks, joints, and laterals—are all dark features that cannot be distinguished by intensity criteria alone.

The goal of our research is to develop an automated method that, given a pipe image, classifies each pixel in the image into one of five classes: background, crack, hole, joint, and lateral. In principle, after the image has been segmented into its classes, each class could be separated further into extents of distress (minor crack, major crack, multiple crack, etc.). In general, an image segmentation and classification problem is difficult to automate because the differences between classes such as joints and cracks, although obvious to a human, can be very difficult to encode mathematically at the pixel level (Sinha, 2000).

##### 4.1 Preprocessing of pipe images

There are several hindrances to identifying the features in underground pipe images. In the case of cracks, the principal difficulty is that cracks are set against a highly patterned background and thus discriminating edges are often not present. A different problem arises when attempting to identify the crack boundary where there is low contrast between the inside of the crack surface and the surrounding area. Pipe background surface is primarily determined by the color of the pipe material. As such, it is usually found in a restricted range of intensities. Typically the gray-scale histogram information is used to enhance the contrast between the object and the background, as discussed in the background chapter. Here it is shown that pattern classification techniques applied to color images can also be used to enhance the contrast between the background surface and the objects present in the pipe image.

**4.1.1 Bayesian classification.** Identification of the boundary between the objects and their surround is formulated as a pattern classification problem. Specifically, it is desired to classify pixels as to whether it is more likely that they came from the objects or the neighboring background regions. In a Bayesian framework, a color pixel  $x_c = (r, g, b)^T$  can be classified as a crack if its *a posteriori* probability  $P(\text{Crack} | x)$  is greater than the corresponding *a posteriori* probability for the surrounding pipe background  $P(\text{Back} | x)$ . If the class-conditional probability densities  $p(x | \text{Crack})$  and  $p(x | \text{Back})$  are known, or can be learned from training images, then Bayes' rule can be used to compute the corresponding *a posteriori* probabilities.

Standard parametric or nonparametric techniques can be used to learn the underlying class-conditional

densities  $p(x|Crack)$  and  $p(x|Back)$ . However, one must bear in mind that the *a posteriori* probabilities  $P(Crack|x)$  and  $P(Back|x)$  are evaluated for each pixel, along each search line, at each time step. Thus, in order for this approach to be usable in practical (real-time) systems, a premium is placed on the online processing time required to discriminate between the classes. Toward this end, Fisher's linear discriminant (Duda and Hart, 1970) is used to enhance the contrast between the objects and the pipe background.

**4.1.2 Fisher's linear discriminant.** Since we intend to apply morphological operators, we require a gray-scale image, in which each pixel  $x_g$  is a scalar  $x_g = w^T x_c$ , where  $w$  is some linear projection. In the case of a two-class discrimination problem, such as distinguishing between cracks and pipe background, Fisher's linear discriminant (Duda and Hart, 1970) can be used to determine the axis,  $w$ , onto which vector color data can be projected which preserves as much of the discriminating capability of the color information as possible. The resulting "Fisher linear discriminant" maximizes the separability of the two classes. Crack images representative of those likely to be encountered during scanning of underground pipes can be used to learn the Fisher discriminant axis using the following algorithm.

1. Calculate mean color in class  $k = 1, 2$

$$m_k = \frac{1}{n_k} \sum_{x \in \chi_k} x \quad (5)$$

2. Determine the within class scatter matrices,  $k = 1, 2$

$$S_k = \sum_{x \in \chi_k} (x - m_k)(x - m_k)^T \quad (6)$$

3. Find the Fisher discriminant vector

$$w = S_w^{-1}(m_1 - m_2) \quad \text{where} \quad S_w = S_1 + S_2 \quad (7)$$

Pipe images representative of those likely to be encountered during object recognition and classification can be used to learn the Fisher discriminant axis. Figure 9 shows that Fisher's discriminant analysis can be used to enhance the contrast between the pipe background and cracks. In the gray-scale image (Figure 9b) there is little contrast between the background and cracks. The histogram equalization, HIS, and YIQ color models can also be used to provide additional contrast as shown in Figures 9c, 9d, and 9e, respectively; although the pipe image shows high contrast, the crack boundary is blurred and the image is noisy. The projection onto the Fisher axis (Figure 9f) enhances the contrast and enables better extraction of the crack features.

## 4.2 Segmentation of pipe images

In underground pipe image segmentation, the following classes are of general interest: the pipe joints (horizontal dark straight lines), pipe laterals (circular dark objects), surface cracks (irregularly shaped thin dark lines), and the pipe background (anywhere from a smooth to a highly patterned surface) (Sinha et al., 1999). The goal of our research is to segment pipe joints, laterals, and cracks based on the *geometric* differences between them, specifically based on morphological techniques (Serra, 1986).

We performed a morphological opening operation on the underground pipe image with increasing sizes of the circular and horizontal structuring elements. Clearly, as the size of the structuring element is increased, features of increasing size are removed by the morphological opening. For example, a structuring element of intermediate size will preserve laterals and a collapsed pipe but will remove cracks and small holes. Figure 10 plots the average area of objects in each class (crack, hole, joint, etc.) based on circular structuring element. That is, if we let  $|t(I)|$  represent the number of dark pixels in  $I$  after binary thresholding, then Figure 10 actually plots the normalized areas

$$a_L(r) = \frac{|t(I \cdot S_C(r))|}{|t(I_L)|} \quad (8)$$

$$a_J(l) = \frac{|t(I \cdot S_H(l))|}{|t(I_J)|} \quad (9)$$

where  $I_L, I_J$  are idealized, prototype images of the perfect lateral and joint.

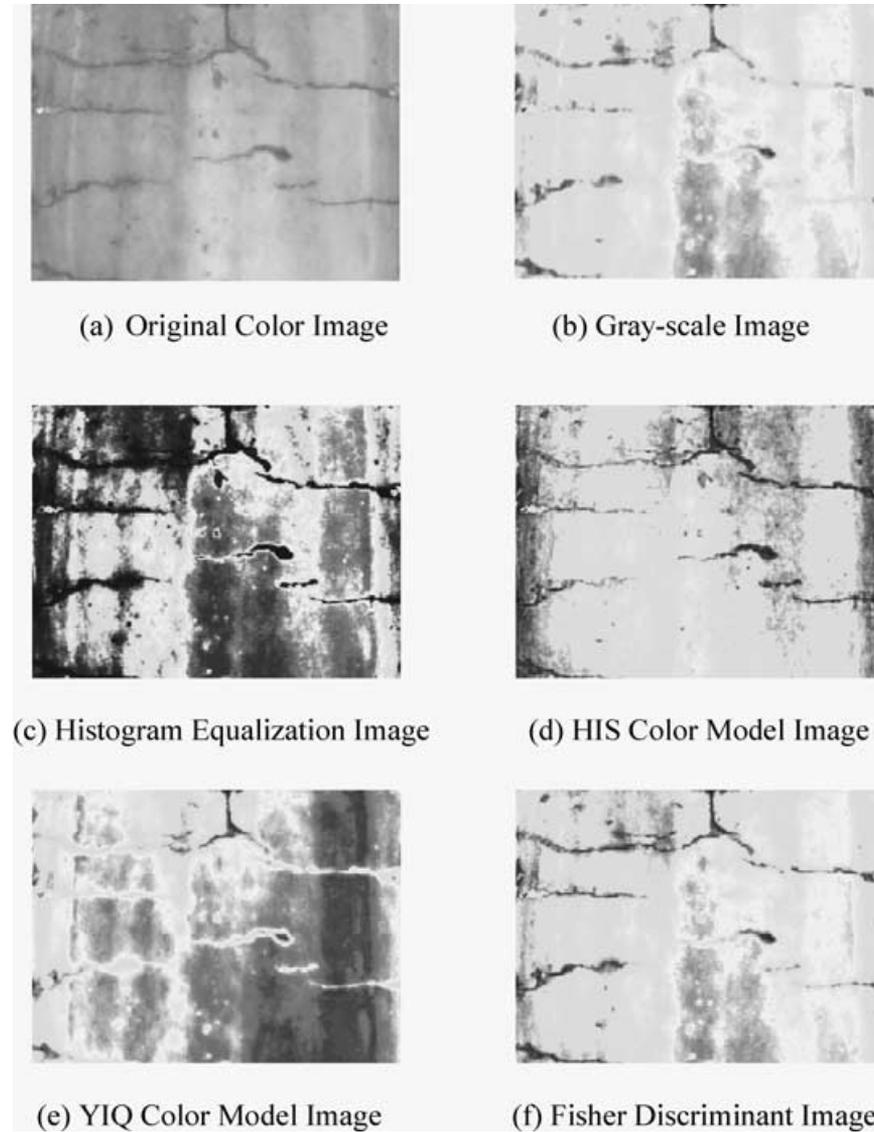
Although the plots in Figure 10 are interesting and intuitive, in order to accurately isolate and classify different objects in an image we have to take into account the *variations* in the area of each class. That is, holes, laterals, etc. all come in a range of sizes, and this range *must* be taken into account in selecting the appropriate structuring element to serve as a classifier. We can compute or assess the ability of any structuring element to discriminate between any two classes (e.g., crack and hole) by examining the degree to which the two classes are separated relative to their standard deviations:

$$D_{i,j}(r) = \frac{|\mu_i(r) - \mu_j(r)|^2}{\sigma_i^2(r) + \sigma_j^2(r)} \quad (10)$$

$$\mu_i(r) = \langle a_L(r) \rangle_i \quad (11)$$

$$\sigma_i^2(r) = \langle a_L(r)^2 \rangle_i - \langle a_L(r) \rangle_i^2 \quad (12)$$

where  $\langle \cdot \rangle_i$  represents an average taken over images of class  $i$ . A parallel definition exists for discriminant  $D_{i,j}(l)$  based on a horizontal structuring element. The value of  $r$  for which  $D_{i,j}(r)$  is maximized represents the optimal feature by which to discriminate between classes  $i$  and  $j$



**Fig. 9.** Fisher's discriminant analysis can be used to enhance the contrast between the pipe background and cracks. In gray-scale image (b) there is little contrast between the background and cracks. The histogram equalization (c) enhances the contrast of the image by transforming the values in an intensity image. The HIS and YIQ color models can also be used to provide additional contrast, as in (d) and (e), respectively; although the pipe image shows high contrast, the crack boundary is blurred and the image is noisy. Projection onto a Fisher axis (f) enhances the contrast and enables better extraction of the crack features.

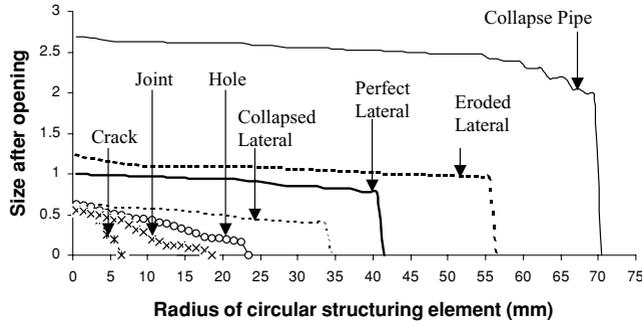
on the basis of the area (i.e., number of pixels) remaining after a morphological opening by element  $S_C(r)$  as shown in Figure 11. By plotting  $D_{i,j}(r)$  and  $D_{i,j}(l)$  for different classes  $i, j$  we can deduce the set of features to be extracted for classification.

More than 100 images, mostly underground concrete pipe from 15 major cities in North America, are used in the proposed morphological segmentation algorithm in a series of steps to isolate the pipe joints and laterals. Figure 12 shows the capability of morphological opening

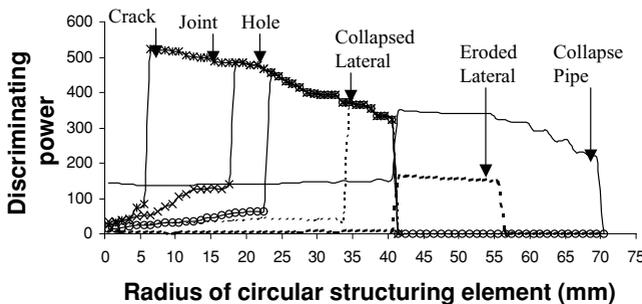
operators for extraction of pipe lateral and joint from underground pipe scanned images.

### 4.3 Detection of defects in pipe images

In the computer vision literature, one can find various techniques addressing different types of data, including natural and artificial textures, synthetic aperture radar images, and magnetic resonance images. In analyzing underground pipe scanned image data, one needs to



**Fig. 10.** Morphological analysis based on a circular structuring element: the average area of each class is plotted as a function of the structuring element diameter; area is normalized to that of an ideal lateral. Clearly, as the diameter is increased, classes with thin, elongated geometries (e.g., cracks, joints) are quickly eliminated.



**Fig. 11.** Lateral discrimination: we can distinguish between laterals and other classes by opening with a circular structuring element, as in Figure 10. We can plot the ability to discriminate ' $D$ ' between an ideal lateral and any other class as the difference in response (normalized to standard error) to the structuring element.

consider complications due to the inherent noise in the scanning process, irregularly shaped cracks, as well as the wide range of pipe background patterns. One of the major problems is detecting defects (especially cracks) that are camouflaged in the background of corroded areas, debris, patches of repair work, and areas of poorly illuminated conditions.

In the past 20 years, many approaches have been developed to deal with the detection of linear features on optic (Geman and Jedynek, 1996) or radar (Samadani and Vesechy, 1990) images. Most of them combine two criteria: (1) a local criterion evaluating the radiometry on some small neighborhood surrounding a target pixel to discriminate lines from background, and (2) a global criterion introducing some large-scale knowledge about the structures to be detected.

Concerning the local criterion, most of the techniques used for pavement distress detection in scanned images

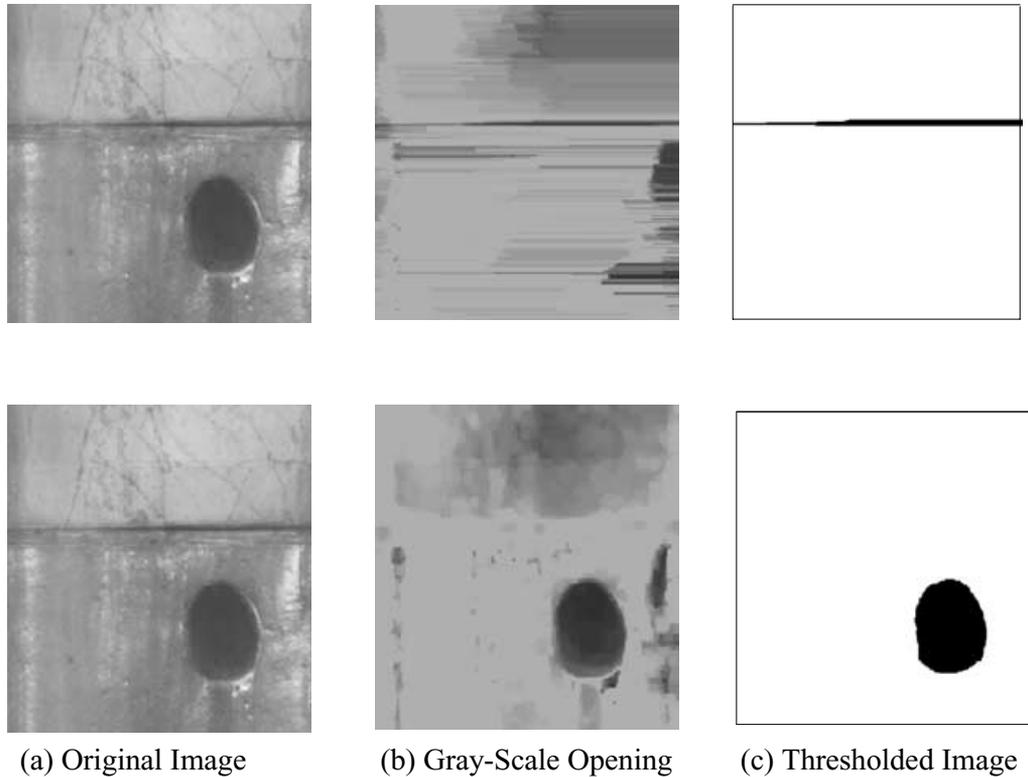
are based either on conventional edge or line detectors (Mohajeri and Manning, 1991; Walker and Harris, 1991). These methods evaluate differences of averages, implying noisy results and variable false-alarm rates (Koutsopoulos and Downey, 1993). In addition, local criteria are in many cases insufficient for edge or line detection, and global constraints must be introduced (Tupin et al., 1998). For instance, dynamic programming is used to minimize some global cost functions, as in the original algorithm of Fischler et al. (1981) and its improvement (Merlet and Zerubia, 1996). Hough transform-based approaches have also been tested for the detection of parametric curves, such as straight lines or circles (Skingley and Rye, 1987). Tracking methods are another possibility. They find the minimum cost path in a graph by using some heuristics, for instance, an entropy criterion (Skingley and Rye, 1987). Energy minimization curves, such as snakes, have been applied (Skingley and Rye, 1987). The Bayesian framework, which is well adapted for taking some contextual knowledge into account, has been widely used (Tupin et al., 1998).

The approach proposed in this paper falls within the scope of the Bayesian framework. Since our aim is to detect the defects present in an image, contextual knowledge on the scale of pixels is insufficient and results in numerous small, disconnected segments. However, on the scale of segments, *a priori* knowledge allows for the detection of cracks. Thus, detection of cracks is performed in two steps. In the first step, crack-segment candidates are detected. In the second step, cracks are obtained by cleaning and linking operations.

The algorithm begins with performing a local detection of cracks (Fieguth and Sinha, 1999). This is based on the fusion of the results from two crack detectors D1 and D2, both taking the statistical properties of image into account. Crack detector D1 is based on a ratio edge detector for which an in-depth statistical study of its behavior is given in Lopes et al. (1993). Detector D2, which has emerged from this research, uses the normalized centered correlation between two populations of pixels. Both responses from D1 and D2 are merged to obtain a unique response as well as an associated direction in each pixel. The detection results are postprocessed to provide candidate segments. Figure 13 shows the different steps of the proposed crack detection algorithms.

In this article, an almost unsupervised method has been proposed for detecting the cracks, as seen in underground pipe scanned images. The method includes both high- and low-level treatments. All the parameters for the detection of cracks are determined experimentally. Thresholded responses of the crack detectors after fusing and linking operations are shown in Figure 14.

The first image (left panel of Figure 14) is a part of Toronto sewer pipeline system, showing some minor

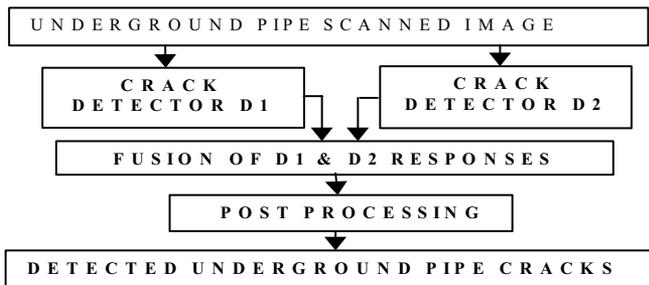


(a) Original Image

(b) Gray-Scale Opening

(c) Thresholded Image

**Fig. 12.** This figure illustrates joint/lateral discrimination using different structuring elements: a horizontal element (top) of length 285 mm, consistent with the geometry of a perfect joint, as opposed to a circular element (bottom) of radius 57 mm, tuned to the shape of a perfect lateral.



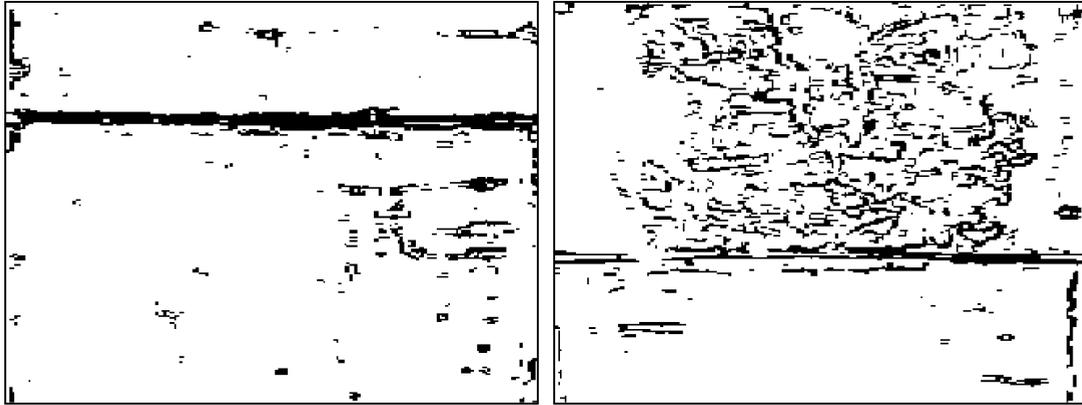
**Fig. 13.** Flowchart showing proposed steps of crack detection algorithm.

cracks in the pipe surface. In this case, the crack detection step performs quite well in detecting most of the crack structures in the image. The second scanned image (right panel of Figure 14) is from the city of Boston. This image has dark background pipe surface with multiple cracks, most severely above the pipe joint. In this case, the crack detection step performed well, but results are noisy with many false alarms. The cleaning and linking operations proved to be a powerful method to fill gaps between the detected segments providing a map of the cracked pipe

surface, while suppressing most of the false-alarm detection. In fact, the results are close to those that could be obtained by a trained human operator for classification of defects in the underground pipes.

#### 4.4 Feature extraction of pipe defects

Feature extraction is an important stage for any pattern recognition task, especially for pipe defect classification, since pipe defects are highly variable and it is difficult to find reliable and robust features. According to the study in Hogg (1993), trained operators mainly rely on five criteria in visual interpretation of images. These are intensity, texture, size, shape, and organization. The intensity corresponds to the spectral features, which can generally be extracted easily. Textural features are those characteristics such as smoothness, fineness, coarseness, or a particular pattern associated with an image (Haralick, 1973). They reflect the local spatial distribution property in a certain region. The spectral and textural features are most widely used in automatic object classification. Other features such as size, shape, and organization information attribute to the large-scale or global spatial distribution.



**Fig. 14.** Thresholded responses of the crack detectors for minor cracks (left) and multiple cracks (right).

Generally, two broad categories of object features are most commonly used in the material/pavement classification field (Kaseko and Ritchie, 1993): shape and textural features. The first class of features, which plays a more important role for object classification, extracts the information based on the geometric shape of the object. Some of the most commonly used methods in this category include area, length, roundness, etc. The second categories (i.e., textural features) distinguish objects by using statistical measures based on gray-scale co-occurrence matrix (Haralick, 1973) and its variant, such as gray-scale difference vector, moment invariants, and gray-scale difference matrix. The salient features of the data can also be extracted through a mapping, such as Fourier transform, discrete cosine transform, Karhunen-Loeve transform, or principal component method (Jain, 1989), from a higher dimensional input space to a lower dimensional representation space.

Depending on the analyzed parameters or features of each object, the most suitable set of features that represents the characteristics of each object in the underground pipe images is selected. We have used information based on the geometric shape and size of the objects present in the underground pipe images for feature extraction. The advantages of the proposed extraction of geometrical features from the image are its capability to quantify distress features in terms of understanding parameters (area, lengths, roundness, etc.) and its ability to classify the segmented image based on such quantities. These features constitute the input parameters for the classifier.

In the present case for classification of severity of cracks, if the attributes are selected to be the major/minor axis length, and area, then the classifier can be trained for classifying different objects based on their geometry. For example, if an object has a width (minor axis length) of a few millimeters and its length (major

axis length) is much greater than its width, then the object can be classified as a crack. On the other hand, if the ratio of major and minor is close to one and its minor axis length is a few centimeters, then the object can be classified as a hole rather than a crack. We can also use the information of four direction projections to classify the cracks. An image projection in a direction is done by adding the pixels values along four directions: horizontal ( $0^\circ$ ), vertical ( $90^\circ$ ), and two diagonals ( $45^\circ$  and  $135^\circ$ ). The idea of using image projections to classify crack types is based on: (1) if it is a longitudinal crack, there is a peak in the vertical projection; (2) if it is a transverse crack, there should be a peak in the horizontal projection; (3) if the crack is diagonal, there is a peak in the diagonal direction; and (4) if it is a mushroom crack, there are peaks in all four direction projections. The 12 features selected for classification of the type of the cracks in the underground pipe image are as follows:

1. Area
2. Number of objects
3. Major axis length
4. Minor axis length
5. Projection of pixels and then taking mean and variance in each of the four projected directions ( $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ , and  $135^\circ$ )

Each segmented crack image is to be classified into one of the following five classes. They are based on the extracted 12 feature vectors, which describe the existence of crack segments and severity of holes present in the image:

1. Minor crack
2. Major crack
3. Minor hole
4. Major hole
5. Fracture pipe

#### 4.5 Classification of pipe defects

Underground pipe defects appear in the form of randomly shaped cracks. The decision making of the pipe condition by human experts is based on very complicated rules such as “if the total area of crack is  $A$ , then it gives a penalty  $f$  to the decision; if the total area of crack is  $B$ , then it gives a penalty  $g$  to the decision; if a pipe has  $f, g, \dots, k$  penalties, then the final decision of the pipe is  $P^{\text{th}}$  class.” To set all these complicated rules, many efforts and time-consuming discussions would be required by human experts. In practice, carrying out this task would be even harder if different criteria existing among the experts about the defects were taken into account. Therefore, there has been a lack of normalization in assessment of underground pipe condition.

For such a complicated decision rule problem, the solution is based on the use of a neural network paradigm that can mimic the human reasoning (Hsieh, 1993). The benefit of the neural network is the generalization ability (Lawrence, 1991) about the untrained samples due to the massively parallel interconnections and easiness of implementation for any complicated rule or mapping problem.

**4.5.1 Error backpropagation algorithm.** In this section crack classification using an error backpropagation (EBP) algorithm (Rumelhart and McClelland, 1986) is discussed. The conventional EBP algorithm used a fixed learning rate and momentum factor; thus, to reduce the learning time and to avoid local minima, these parameters must be determined adaptively. A variable learning rate and momentum factor are used by iteratively updating weights, resulting in the modified EBP algorithm. In the modified EBP algorithm, the learning rate  $\eta$ , momentum factor  $\alpha$ , and weight  $\omega$  are updated by the following equations, respectively:

$$\Delta\eta_{jk}(m+1) = \varepsilon \left( \frac{\partial E_p}{\partial \omega_{jk}} \right)^2 + \beta \Delta\eta_{jk}(m), \quad (13)$$

$$\Delta\alpha_{jk}(m+1) = \mu \left( \frac{\partial E_p}{\partial \omega_{jk}} \right)^2 + \gamma \Delta\alpha_{jk}(m), \quad (14)$$

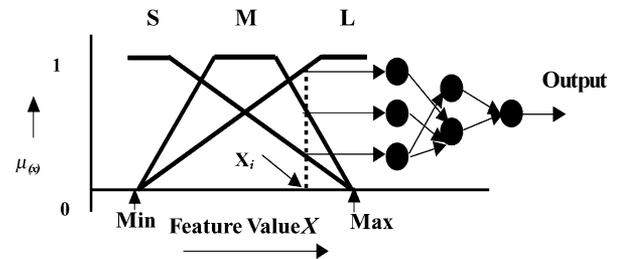
$$\Delta\omega_{jk}(m+1) = -\eta_{jk} \frac{\partial E_p}{\partial \omega_{jk}} + \alpha_{jk} \Delta\omega_{jk}(m) \quad (15)$$

where  $\varepsilon$ ,  $\beta$ ,  $\mu$ , and  $\gamma$  denote constants and  $m$  represents the iteration step. Subscripts  $j$  and  $k$  signify the  $j$ th neuron of the input (hidden) layer and  $k$ th neuron of the hidden (output) layer, respectively.  $E_p$  represents the total error function at the  $p$ th layer. The network is constructed by 12 input neurons, 7 neurons in the hidden layer, and 5 output neurons for five classes of pipe defects. Parameter values are selected experimentally: in the conventional

algorithm, the learning rate and momentum factor are set to 0.7 and 0.15, respectively, whereas in the modified algorithm, 0.5 and 0.2, respectively.

**4.5.2 Classification using neuro-fuzzy algorithm.** To increase the recognition rate, a neuro-fuzzy algorithm is employed that combines neural networks and the fuzzy concepts. Neural networks have learning capability and the fuzzy concepts can absorb variability in feature values. The fuzzy concept (Zadeh, 1965) can be combined with neural networks in various ways (Pal and Mitra, 1992; Glorennec, 1994). In this study the fuzzy concept is applied simply in converting feature values into fuzzified data, which are inputs to the modified backpropagation neural network algorithm. In the proposed neuro-fuzzy algorithm, the fuzzy data are used as inputs to neural networks. Sometimes variation of feature values is large, and then it is difficult to classify defects correctly based on these feature values. To solve this problem, each defect feature value is first converted into three fuzzy data (Yamakawa and Teodorescu, 1997), and then learning is performed with these 3I fuzzy data using the modified EBP algorithm. Finally, defects are classified using the modified backpropagation algorithm.

To convert 12 normalized features into 36 fuzzy data, the *MAX* and *MIN* values are determined that are the maximum and minimum feature values for the entire data set, respectively. As shown in Figure 15, three membership functions denoted by “S” (small), “M” (medium), and “L” (large) are generated. Note that these membership functions are specified by *MIN* and *MAX*, as shown in Figure 15. Then three fuzzy data are computed for each feature values and these data are used as the input data to neural networks. In Figure 15  $\mu_S(x_i)$ ,  $\mu_M(x_i)$ , and  $\mu_L(x_i)$  are three fuzzy data of an input feature value ( $x_i$ ), corresponding to linguistic variables of “S,” “M,” and “L,” respectively. The trapezoidal membership function, as shown in Figure 15, is located at the average value of the features of the same defect, and it has a maximum value of 1 over the limited range that is specified by the standard deviation of the feature value.



**Fig. 15.** Proposed neuro-fuzzy network for classification of underground pipe images.

**Table 1**

Classification rate for the proposed neuro-fuzzy network

<i>Classification method</i>	<i>Classification rate (%)</i>
Euclidean distance method	81.1
Fuzzy K-nearest neighbor method	83.2
Backpropagation neural network	85.9
Proposed neuro-fuzzy network	89.6

To generate a linguistic variable, the average and standard deviation of the feature values of the defect are computed. Then the interval between *MIN* and *MAX* is uniformly divided into several subintervals, where *MIN* and *MAX* represent the minimum and maximum of average values of the specific feature, respectively. The membership function of each image is centered at the average value of the features of the defect. Variation of feature values for the same image is allowed by employing the trapezoidal membership function (i.e., the width at the top of the trapezoidal membership function is set to  $\sigma_i$ , where  $\sigma_i$  denotes the standard deviation of the *i*th feature value).

**4.5.3 Experimental results.** For performance comparison with the proposed neuro-fuzzy algorithm, conventional algorithms such as Euclidean distance method, EBP algorithm, and fuzzy methods with triangular and trapezoidal membership functions are simulated. In the modified EBP algorithm to reduce the computational complexity and to avoid local minima problem, the learning rate and momentum factor are varied adaptively. In the input layer, there exist 12 nodes for 12 features, but 5 neurons in the output layer for 5 classes of pipe defects. The number of nodes in the hidden layer is determined experimentally. In the proposed neuro-fuzzy algorithm 36 fuzzy data are used in the input layer, and in the hidden and output layer the same number of neurons are used as in the modified EBP algorithm. Table 1 shows the classification results by the proposed neuro-fuzzy algorithm and other conventional algorithms. From Table 1 it can be observed that the proposed neuro-fuzzy algorithm using a trapezoidal membership function yields better classification results than the Euclidean distance method, EBP algorithm, and fuzzy-based algorithms.

## 5 CONCLUSIONS

A practical use of computer vision for automatic segmentation of underground concrete pipes from PSET scanned images has been suggested and implemented. Two main obstacles are the poor background illumina-

tion and the highly patterned surface of sewer pipes, which causes problems for detection of defects. The defect detection techniques used in this study, especially the linear feature extraction, are robust for detection of cracks, but for root extraction, the techniques need to be improved further. The local crack detector deals with scanned images with respect to their statistical properties. Since there does not appear to exist a single coherent model suitable for reliable detection of pipe cracks, it is essential that some means of integrating information from multiple image operators and knowledge sources be devised. This research has provided a simple mechanism for integrating the information provided by the two operators for the specific task of crack detection. For joint and lateral analysis, a technique based on mathematical morphology has been proposed and investigated. The technique is quite effective if the original size of the joint and lateral is available or estimated accurately.

A fully automated underground pipe inspection system is envisaged, in which successive image frames are analyzed until a pipe defect is identified. The defective frame would be selected and analyzed using the techniques discussed in this study. Such a scheme would extract valuable information from PSET surveys of underground concrete pipes. The proposed automated system has the potential to overcome the limitations of the current CCTV inspection and can provide a more accurate assessment of underground sewer pipe conditions.

## ACKNOWLEDGMENTS

This research is supported by the Natural Science and Engineering Research Council (NSERC) of Canada. The authors would like to thank CORE Corp., California, for providing the scanned images, and the municipalities and consultants in North America for their valuable input.

## REFERENCES

- Agin, G. J. (1980), Computer vision systems for industrial inspection and assembly, *Computer*, **13**, 11–20.
- Aydin, T., Yemez, Y., Anarim, E. & Sankur, B. (1996), Multidirectional and multiscale edge detection via M-band wavelet transform, *IEEE Transactions on Image Processing*, **5**, 1370–77.
- Ballard, D. H. & Brown, C. M. (1982), *Computer Vision*, Prentice-Hall, Englewood Cliffs, NJ.
- Bassim, M. N. & Houssny-Emam, M. (1983), Acoustic emission, In: *Time and Frequency Analysis of Acoustic Emission Signals*, Gordon and Breach Scientific Publishers, Newark, NJ, pp. 139–63.
- Birks, A. & Green, R. (1991), *Nondestructive Testing Handbook*, 2nd ed., vol. 7, American Society for Nondestructive Testing, Columbus, OH.

- Castleman, K. R. (1996), *Digital Image Processing*, Prentice-Hall, Inc., Englewood Cliffs, NJ.
- Chang, Y. L. & Li, X. (1995), Fast image region growing, *Image and Vision Computing*, **13**, 559–71.
- Chin, R. T. & Harlow, C. A. (1982), Automated visual inspection: a survey, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **4**, 557–73.
- Chow, C. K. & Kaneko, T. (1972), Automatic boundary detection of the left ventricle from cineangiograms, *Computers in Biomedical Research*, **5**, 388–410.
- Davies, T. G. & Mamlouk, M. S. (1985), Theoretical response of multilayer pavement system to dynamic non-destructive testing, *Transportation Research Record*, **1022**, 1–7.
- Demigny, D., Lorca, F. G. & Kessal, L. (1995), Evaluation of edge detectors' performances with a discrete expression of Canny's criteria, *Proceedings of International Conference on Image Processing*, IEEE, Los Alamitos, CA, pp. 169–72.
- Duda, R. O. & Hart, P. E. (1970), *Pattern Classification and Scene Analysis*, Wiley Publications, New York, pp. 271–2.
- Fieguth, P. W. & Sinha, S. K. (1999), Automated analysis and detection of cracks in underground scanned pipes, *IEEE Image Processing Conference*, Kobe, Japan, pp. 395–9.
- Fischler, M. A., Tevenbaum, J. M. & Wolf, H. C. (1981), Detection of roads and linear structures in low resolution aerial imagery using multisource knowledge integration technique, *Computer Graphics Image Processing*, **15**(3), 201–23.
- Fu, K. S. (1982), *Syntactic Pattern Recognition and Applications*, Prentice-Hall, Englewood Cliffs, NJ.
- Geman, D. & Jedynak, B. (1996), An active testing model for tracking roads in satellite images, *IEEE Transactions on Pattern Analysis of Machine Intelligence*, **18**, 1–14.
- Glorennec, P. Y. (1994), *Learning Algorithms for Neuro-Fuzzy Networks*, Fuzzy Control Systems, Kandel A., and Langholz, G. (eds.), CRC Press, Boca Raton, FL.
- Gonzalez, R. C. & Wintz, P. (1987), *Digital Image Processing*, Addison-Wesley Publishing Co., Reading, MA.
- Haralick, R. M. (1973), Textural features for image classification, *IEEE Transactions on Systems, Man, and Cybernetics*, **3**, 610–21.
- Haralick, R. M. & Shapiro, L. G. (1985), Image segmentation techniques, *Computer Vision, Graphics, and Image Processing*, **29**, 100–32.
- Hibino, Y., Nomura, T., Ohta, S. & Yoshida, N. (1994), Laser scanner for tunnel inspections, *Journal of International Water Power and Dam Construction*, **6**, 17–28.
- Hogg, D. C. (1993), Shape in machine vision, *Image and Vision Computing*, **11**, 309–16.
- Hsieh, C. (1993), Some potential applications of artificial neural systems in financial management, *Journal of Systems Management*, April, 12–15.
- Iseley, T. (1999), Pipeline condition assessment: achieving uniform defect ratings, *Proceedings of Infra 99 International Conference*, **1**, 121–32.
- Iseley, T., Abraham, D. M. & Gokhale, S. (1997), Intelligent sewer condition evaluation technologies, *Proceedings of International NO-DIG Conference*, pp. 254–65.
- Jain, A. K. (1989), *Fundamentals of Digital Image Processing*, Prentice-Hall, Englewood Cliffs, NJ.
- Jain, A. K. & Chandrasekaran, B. (1982), *Dimensionality and Sample Size Consideration in Pattern Recognition Practice*, Handbook of Statistics, North Holland Publishing Co., New York.
- Joynson, R. E., McCary, R. O., Oliver, D. W., Silverstein-Hedengren, K. H. & Thumhart, L. L. (1986), Eddy current imaging of surface breaking structures, *IEEE Transactions on Magnetics*, **22**, 11–24.
- Kanal, L. & Chandrasekaran, B. (1971), On dimensionality and sample size in statistical pattern recognition, *Pattern Recognition*, **3**, 225–34.
- Kaseko, M. S. & Ritchie, S. G. (1993), A neural network based methodology for pavement crack detection and classification, *Transportation Research*, **1**(4), Transportation Research Board, Washington, D.C., pp. 275–91.
- Kittler, J. & Illingworth, J. (1985), On threshold selection using clustering criteria, *IEEE Transactions on Systems, Man, and Cybernetics*, **15**(5), 652–5.
- Koutsopoulos, H. N. & Downey, A. B. (1993), Primitive based classification of pavement cracking images, *Journal of Transportation Engineering, ASCE*, **119**(3), 402–18.
- Krstulovic, O. N., Woods, R. D. & Al Shaye, N. (1996), Non-destructive testing of concrete structures using the Rayleigh wave dispersion method, *ACI Materials Journal*, **93**, 75–86.
- Kundu, A. & Mitra, S. K. (1987), A new algorithm for image edge extraction using a statistical classifier approach, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **9**(4), 569–77.
- Law, T., Itoh, H. & Seki, H. (1996), Image filtering, edge detection, and edge tracing using fuzzy reasoning, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **18**, 481–91.
- Lawrence, J. (1991), *Introduction to Neural Networks*, Third Edition, California Scientific Software, Grass Valley, CA.
- Lopes, E., Nezry, R., Touz, R. & Laur, H. (1993), Structure detection, and statistical adaptive filtering in SAR images, *International Journal of Remote Sensing*, **14**(9), 1735–58.
- Merlet, N. & Zerubia, J. (1996), New prospects in line detection by dynamic programming, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **8**, 426–31.
- Mohajeri, M. H. & Manning, P. J. (1991), ARIA: an operating system of pavement distress diagnosis by image processing, *Transportation Research Record*, **1311**, Transportation Research Board, Washington, D.C., pp. 120–30.
- Newman, T. S. & Jain, A. K. (1995), A survey of automated visual inspection, *Computer Vision Image Understanding*, **61**(2), 231–62.
- Oja, E. (1983), *Subspace Methods of Pattern Recognition*, Research Studies Press, Letchworth, England.
- Pal, N. R. & Pal, S. K. (1987), Segmentation based on contrast homogeneity measure and region size, *IEEE Transactions on Systems, Man, and Cybernetics*, **17**(5), 857–68.
- Pal, S. K. & Mitra, S. (1992), Multilayer perceptron, fuzzy sets, and classification, *IEEE Transactions on Neural Networks*, **3**(5), 683–97.
- Pratt, W. K. (1978), *Digital Image Processing*, Wiley, New York.
- Priese, L. & Rehrmann, V. (1993), On hierarchical color segmentation and application, *Proceedings in Computer Vision and Pattern Recognition*, Los Alamitos, CA, 633–4.
- Ramesh, V. & Haralick, R. M. (1994), An integrated gradient edge detector: theory and performance valuation, *Proceedings in ARPA Image Understanding Workshop*, Monterey, CA, ARPA, Los Altos, CA, 689–702.
- Rosenfeld, A. & Kak, A. C. (1982), *Digital Picture Processing*, Academic Press, New York, 2nd edition.
- Rumelhart, D. E. & McClelland, J. (1986), *Parallel Distributed Processing*, **1**, MIT Press, Cambridge, MA.
- Samadani, R. & Vesechy, J. F. (1990), Finding curvilinear features in speckled images, *IEEE Transactions on Geoscience Remote Sensing*, **28**, 669–73.

- Schalkoff, R. J. (1992), *Pattern Recognition: Statistical, Structural, and Neural Approaches*, John Wiley and Sons, Inc., New York.
- Schettini, R. (1993), A segmentation algorithm for color images, *Pattern Recognition Letters*, **14**, 499–506.
- Serra, J. (1986), Introduction to mathematical morphology, *Journal of Computer Visual Graphics Image Processing*, **35**, 283–305.
- Shaikh, M. A. & Tian, B. (1996), Neural network based cloud detection/classification using textural and spectral features, *Proceedings of IGRASS*, Lincoln, NE, 1105–7.
- Sinha, S. K. (2000), *Automated Underground Pipe Inspection Using a Unified Image Processing and Artificial Intelligence Methodology*, Ph.D. Thesis, University of Waterloo, Ontario, Canada.
- Sinha, S. K., Fieguth, P. W., Polak, M. A. & McKim, R. A. (1999), Automated underground pipe condition assessment by image analysis of the state-of-the-art sewer scanner and evaluation technology surveys, *Proceedings North American No-Dig'99 Conference*, Orlando, FL, 63–77.
- Skingley, J. & Rye, A. J. (1987), The Hough transform applied to images for thin line detection, *Pattern Recognition Letters*, **6**, 61–7.
- Smith, A. R. (1978), Color gamut transform pairs, *Computer Graphics*, **2**(3), 12–19.
- Tupin, F., Maitre, M., Mangin, J., Nicolas, J. & Pechersky, E. (1998), Detection of linear features in SAR images: application to road network extraction, *IEEE Transactions on Geoscience Remote Sensing*, **36**(2), 434–52.
- Van Cauwelaert, F. J., Alexander, D. R., White, T. D. & Barker, W. R. (1989), Multilayer elastic program for backcalculating layer moduli in pavement evaluation, *Nondestructive Testing of Pavements and Backcalculation of Moduli*, ASTM STP 1026, ASTM, pp. 171–88.
- Vrabel, M. J. (1996), Edge detection with a recurrent neural network, *Proceedings of Application and Science of Artificial Neural Networks II*, **2760**, SPIE, Bellingham, WA, 365–71.
- Walker, R. S. & Harris, R. L. (1991), Noncontact pavement crack detection systems, *Transportation Research Record*, **1311**, Transportation Research Board, 149–57.
- Weszka, J. S. & Rosenfeld, A. (1979), Histogram modification for threshold selection, *IEEE Transactions on Systems, Man, and Cybernetics*, **9**(1), 38–52.
- Wirahadikusumah, R., Abraham, D. M., Iseley, T. & Prasanth, R. K. (1998), Assessment technologies for sewer system rehabilitation, *Journal of Automation in Construction*, **7**(4), 259–70.
- Yamakawa, T. & Teodorescu, H. (1997), *Neuro-Fuzzy Systems Hybrid Configuration*, Kluwer Academic Publishers, Boston, MA.
- Zadeh, L. A. (1965), Fuzzy sets, *Information Control*, **8**, 338–53.