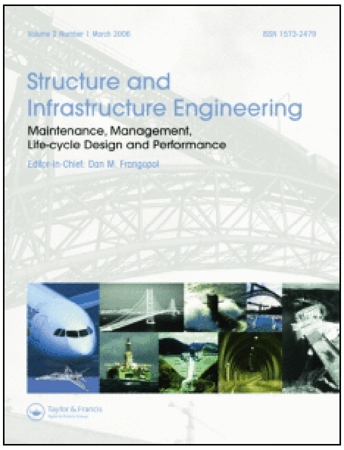


Structure and Infrastructure Engineering
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713683556

A survey and evaluation of promising approaches for automatic image-based defect detection of bridge structures

Mohammad R. Jahanshahi [a]; Jonathan S. Kelly [b]; Sami F. Masri [a]; Gaurav S. Sukhatme [b]
[a] Department of Civil and Environmental Engineering, University of Southern California, Los Angeles, California, USA [b] Computer Science Department, University of Southern California, Los Angeles, California, USA

First Published:December2009

To cite this Article Jahanshahi, Mohammad R., Kelly, Jonathan S., Masri, Sami F. and Sukhatme, Gaurav S.(2009)'A survey and evaluation of promising approaches for automatic image-based defect detection of bridge structures',Structure and Infrastructure Engineering,5:6,455 — 486

To link to this Article: DOI: 10.1080/15732470801945930

URL: http://dx.doi.org/10.1080/15732470801945930

Structure and Infrastructure Engineering
Vol. 5, No. 6, December 2009, 455–486

Taylor & Francis
Taylor & Francis Group

A survey and evaluation of promising approaches for automatic image-based defect detection of bridge structures

Mohammad R. Jahanshahi[a], Jonathan S. Kelly[b], Sami F. Masri[a]* and Gaurav S. Sukhatme[b]

[a]*Department of Civil and Environmental Engineering, University of Southern California, Los Angeles, California, USA;* [b]*Computer Science Department, University of Southern California, Los Angeles, California, USA*

(*Received 4 September 2007; final version received 28 January 2008*)

Automatic health monitoring and maintenance of civil infrastructure systems is a challenging area of research. Nondestructive evaluation techniques, such as digital image processing, are innovative approaches for structural health monitoring. Current structure inspection standards require an inspector to travel to the structure site and visually assess the structure conditions. A less time consuming and inexpensive alternative to current monitoring methods is to use a robotic system that could inspect structures more frequently. Among several possible techniques is the use of optical instrumentation (e.g. digital cameras) that relies on image processing. The feasibility of using image processing techniques to detect deterioration in structures has been acknowledged by leading experts in the field. A survey and evaluation of relevant studies that appear promising and practical for this purpose is presented in this study. Several image processing techniques, including enhancement, noise removal, registration, edge detection, line detection, morphological functions, colour analysis, texture detection, wavelet transform, segmentation, clustering and pattern recognition, are key pieces that could be merged to solve this problem. Missing or deformed structural members, cracks and corrosion are main deterioration measures that are found in structures, and they are the main examples of structural deterioration considered here. This paper provides a survey and an evaluation of some of the promising vision-based approaches for automatic detection of missing (deformed) structural members, cracks and corrosion in civil infrastructure systems. Several examples (based on laboratory studies by the authors) are presented in the paper to illustrate the utility, as well as the limitations, of the leading approaches.

Keywords: image processing; pattern recognition; crack; corrosion; bridge inspection; defect detection

1. Introduction

Change detection by means of digital image processing has been used in several fields, including homeland security and safety (Shinozuka 2003), product quality control (Garcia-Alegre *et al.* 2000), system identification (Shinozuka *et al.* 2001, Chung *et al.* 2004), aircraft skin inspections (Siegel and Gunatilake 1998), video surveillance (Collins *et al.* 2000), aerial sensing (Watanabe *et al.* 1998, Huertas and Nevatia 2000), remote sensing (Goldin and Rudahl 1986, Bruzzone and Serpico 1997, Deer and Eklund 2002, Peng *et al.* 2004), medical applications (Dumskyj *et al.* 1996, Lemieux *et al.* 1998, Thirion and Calmon 1999, Rey *et al.* 2002, Bosc *et al.* 2003), underwater inspections (Lebart *et al.* 2000, Edgington *et al.* 2003), transportation systems (Achler and Trivedi 2004) and nondestructive structural health monitoring (Dudziak *et al.* 1999, Abdel-Qader *et al.* 2003, Sinha *et al.* 2003, Poudel *et al.* 2005). These applications of digital image processing share common steps that have been reviewed by Singh (1989), Coppin and Bauer (1996), Lu *et al.* (2004) and, most recently, by Radke *et al.* (2005). As change detection techniques are problem-oriented, most image processing approaches are limited to detecting only one type of defect at a time. The aim of this review is to present and evaluate the steps and algorithms that are necessary for detecting various changes simultaneously in three-dimensional (3D) truss structures by digital image processing. In this study, the changes of main concern are missing (deformed) parts, cracks and rust; however, other changes (e.g. missing bolts) are also possible to detect using nondestructive techniques such as infrared imaging (Shubinsky 1994).

1.1. Motivation

Traditional bridge inspection is time consuming and expensive because it requires an expert to visually inspect the structure site for changes. Yet, visual inspection remains the most commonly used technique

*Corresponding author. Email: masri@usc.edu

ISSN 1573-2479 print/ISSN 1744-8980 online

DOI: 10.1080/15732470801945930
http://www.informaworld.com

to detect damage, since many of the bridges in the United States are old and not instrumented with sensor systems. In cases of special structures, such as long-span bridges, access to critical locations for visual inspection can be difficult (Pines and Aktan 2002). A robotic system that could inspect structures more frequently and which is accompanied by other nondestructive techniques could be a great advance in infrastructure maintenance. The use of digital cameras, image processing and pattern recognition techniques is an appropriate approach to reach this goal.

1.2. Background

An automatic crack detection procedure in welds based on magnetic particle testing (Coffey 1988) was introduced by Ho *et al.* (1990). This method can only be used on ferromagnetic materials. First, the testing surface is sprayed with white paint to reduce the initial noise of subsequently captured images. Next, a magnetic field is applied to the weld. Then, magnetic ink made of small magnetic particles suspended in oil is sprayed over the testing surface. The change of flux density at the crack causes the magnetic particles to trace out the shape of the crack on the weld surface. Lastly, an image of the prepared surface is captured and cracks are detected by means of the Sobel edge detection operator (Duda and Hart 1973, Gonzalez *et al.* 2004) and by implementing a boundary tracing algorithm. The results were satisfactory as reported by Ho *et al.* (1990), but clearly this technique has drawbacks since a preprocessing step is required.

Tsao *et al.* (1994) composed image analysis and an expert system modulus to detect spalling and transverse cracks in pavements. The overall accuracy of the system for detecting spalling and transverse cracks was reported to be 85% and 90%, respectively (Chae 2001). Kaseko *et al.* (1994) and Wang *et al.* (1998) used the image processing and neural network techniques to detect defects in pavements.

Siegel and Gunatilake (1998) developed a remote visual inspection system of aircraft surfaces. To detect cracks, their proposed algorithm detects rivets as cracks propagate on rivet edges. Multi-scale edge detection is used to detect the edges of small defects at small scales and the edges of large defects at large scales. By tracing edges from high scale to low scale, it is possible to define the propagation depth of edges. Using other features based on wavelet transformation (Prasad and Iyenger 1997, Abtoine *et al.* 2004) and a trained back-propagation neural network (Duda *et al.* 2001), cracks can be classified from other defects such as scratches. Corroded regions can also be detected by defining features based on two-dimensional (2D) discrete wavelet transformation of the captured images and using a neural network classifier (Siegel and Gunatilake 1998).

Nieniewski *et al.* (1999) developed a visual system that could detect cracks in ferrites. A morphological detector based on a top-hat transform (Salembier 1990) detects irregular changes of brightness, which could lead to crack detection. *k*-nearest neighbours (Duda *et al.* 2001) is used as a classifier to classify cracks from grooves. The outcome of this study is very promising, and this technique is quite robust, despite the presence of noise, unlike other edge detection operators used for crack extraction.

Moselhi and Shehab-Eldeen (2000) used image analysis techniques and neural networks to automatically detect and classify defects in sewer pipes. The accuracy rate of the proposed algorithm is 98.2%, as reported by the authors.

Chae (2001) proposed a system consisting of image processing techniques, along with neural networks and fuzzy logic systems for automatic defect (including cracks) detection of sewer pipelines.

Benning *et al.* (2003) used photogrammetry to measure the deformations of reinforced concrete structures. A grid of circular targets is established on the testing surface. Up to three cameras capture images of the surface simultaneously. The relative distances between the centres of adjacent targets make it possible to monitor the evolution of cracks.

Abdel-Qader *et al.* (2003) analysed the efficacy of different edge detection techniques in identifying cracks in concrete pavements of bridges. They concluded that the fast Harr transform (FHT), which is a wavelet transform with a mother wavelet of Harr, has the most accurate crack detection capability in contrast with fast Fourier transforms (FFTs), Sobel and Canny edge detection operators (Bachmann *et al.* 2000, Alageel and Abdel-Qader 2002).

A study on using computer vision techniques for automatic structural assessment of underground pipes has been carried out by Sinha *et al.* (2003). The algorithm proposed by Sinha *et al.* (2003) consists of image processing, segmentation, feature extraction, pattern recognition and a proposed neuro-fuzzy network for classification.

Choi and Kim (2005) applied machine vision techniques to evaluate and classify different surface corrosion damages. They proposed a set of morphological attributes (colour, texture and shape) as appropriate features to be used for classification. Hue-saturation-intensity (HSI) colour space (Gonzalez and Wintz 1987, Pratt 2001) and co-occurrence matrix methods are used for colour and texture features respectively. Principal component analysis (PCA)

(Jolliffe 2002) is used to optimise the selection of the features.

Giakoumis *et al.* (2006) detected the cracks in digitised paintings by thresholding the output of the morphological top-hat transform. Sinha and Fieguth (2006b) detected the defects in underground pipe images by thresholding the morphological opening of the pipe images using different structuring elements.

1.3. Scope

As noted above, the aim of this study is to present and also evaluate the feasibility of the steps that are essential for vision-based automatic health monitoring of structures. The focus of this study is mainly on visual imaging and image processing techniques that can detect crack and corrosion. The ultimate objective is to develop a robotic system that can independently navigate underneath bridges and capture images. Different image acquisition systems, including digital cameras and infrared imaging devices, can be used for this purpose. The system should detect and classify cracks, corrosion, missing parts and deformed members. Apart from detecting defects, the system also should be able to localise the position of the deterioration. Clearly, detecting a defect in a large bridge that contains repeats of the 3D truss system pattern is difficult, even for a human being. Therefore, localisation is crucial to resolving the problem. A set of steps and algorithms that are most promising for reaching this goal is presented in this study.

A quick review of preprocessing techniques is presented in §2, since preprocessing of the captured images might be a prerequisite for the application of other algorithms. Image registration and its role in the defined problem are mentioned in §3. This section is useful for detecting the missing or deformed structural members. In §4, a brief review of pattern recognition concepts is presented, and supervised and unsupervised classification algorithms are introduced. Classification plays an important role in differentiating defects from non-defective changes. Some important and useful classification techniques are discussed in this section. For corrosion detection, the wavelet filter bank approach, which has applications in both crack extraction and texture segmentation, is used. Section 5 briefly reviews wavelet decomposition and reconstruction of images. Section 6 focuses on common edge detection techniques and some morphological techniques useful for crack extraction. The introduction to morphological techniques is included in §6.2. Section 7 discusses promising texture segmentation techniques and colour characteristics of corrosion, which could be a key tool in clustering defective parts from non-defective ones. Finally, conclusions are discussed in §8.

2. Preprocessing

Preprocessing consists of a series of steps that prepare the image for further processing. These enhancement techniques, including image smoothing, image sharpening, contrast modification and histogram modification, can be found in almost any digital image processing book (e.g. Gonzalez and Wintz 1987, Gonzalez and Woods 1992, Pratt 2001, Gonzalez *et al.* 2004).

The purpose of image smoothing is to reduce noise in an image. Below, some practical and useful image smoothing techniques are mentioned.

2.1. Neighbourhood averaging (mean filter)

The average grey-level value of a neighbourhood is replaced as the new value in the smoothed image. Although this technique is very simple, it will blur any sharp edges. To overcome this shortcoming, it is necessary to average the brightness values of only those pixels in the neighbourhood that have similar brightness as the pixel that is being processed. The most important factor in this technique is the neighbourhood and the assigned weights for averaging the values within the neighbourhood. In fact, it is possible to write neighbourhood averaging as a 2D convolution by sliding a kernel over the grey-scale image. This convolution could be written mathematically as:

$$Q(i,j) = \sum_{k=-m}^{m'} \sum_{l=-n}^{n'} I(i+k-1, j+l-1) \times K(k+m+1, l+n+1), \quad (1)$$

where I is the grey-scale $(M+m+m') \times (N+n+n')$ image derived from the initial $M \times N$ image by mirroring the border elements to create a larger matrix. i and j are the pixel coordinates in the convolved image and k and l are parameters used in the summation operator to specify the coordinates of each kernel value in the kernel window. In that way, Q will be $M \times N$ as well. Each value in the image matrix is the brightness value of the relevant pixel and K is the kernel with $m+m'+1$ rows and $n+n'+1$ columns.

Gaussian smoothing kernel can be used as a suitable neighbouring average kernel. This kernel can be estimated as a 2D isotropic Gaussian distribution, as shown in Equation (2):

$$G(x,y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right), \quad (2)$$

where σ is the standard deviation. Since the digital image is a set of discrete pixels, a discrete approximation of the Gaussian distribution is needed. One can assume that the value of the Gaussian distribution for points further than three standard deviations from the mean is zero. Horn (1986) has introduced a way to approximate continuous Gaussian functions with discrete filters.

2.2. *Median filter*

It is possible to overcome the image blurring of the neighbourhood averaging method by choosing a threshold; however, the threshold is usually based on extensive trial and error. For this reason, the grey level of each pixel can be replaced by the median value of the grey level of the neighbouring pixels. This simple nonlinear filter is referred to as a median filter. This technique is suitable when crack detection is of interest.

2.3. *Averaging of multiple images*

Suppose that a noisy image J_k (i, j) is formed as follows:

$$J_k(i,j) = I(i,j) + \eta_k(i,j), \tag{3}$$

where I (i, j) is the original image and η_k (i, j) is noise. Assuming that noise is uncorrelated in different pixels and has zero mean value the average of multiple images can be written as:

$$\overline{J}(i,j) = \frac{1}{N}\sum_{k=1}^{N} J_k(i,j). \tag{4}$$

It then follows that:

$$E\big(\overline{J}(i,j)\big) = I(i,j), \tag{5}$$

and

$$\sigma_{\overline{J}(i,j)} = \frac{1}{\sqrt{N}}\sigma_\eta(i,j), \tag{6}$$

where E is the mathematical expectation (mean) and σ is the standard deviation. Equations (5) and (6) show that, as the number of multiple images increases, the averaged image converges to the original image, and the deviation of the pixel values decreases. This is a suitable way to remove noise; however, the noisy images should be properly registered prior to averaging (Gonzalez and Wintz 1987, Gonzalez and Woods 1992).

3. Image registration

Image registration is the process of matching two or more images of the same scene. These images can be captured at different times, from different orientations, or even by different types of sensors. Image registration is a fundamental task of many change detection procedures. A simple example is to register two images taken at different times and from different orientations, and identify the differences between the latter image and the former reference image as a measure of possible changes. Subtracting the two images easily detects changes; however, the registration of images is not always an easy task, especially if the images include 3D scenes with occlusions. Two comprehensive survey papers on image registration are published by Brown (1992) and Zitová and Flusser (2003).

Image registration could be key to solving problems such as visual detection of missing or deformed structural members. It could also help localise detected changes by 3D reconstruction of truss systems. The crack and corrosion detection techniques discussed in §6 and §7 are image registration-independent because the current image registration algorithms are insufficient for 3D reconstruction of a truss in the presence of occlusions. The authors believe that image registration is a key piece of the puzzle to solve the problem, and for this reason a quick review is presented here. Most image registration techniques consist of feature detection, feature matching, transformation and image resampling, as described below.

3.1. *Feature detection*

In feature detection, control points such as distinctive objects, edges, topographies, points, line intersections and corners are detected. From a robotic standpoint, it is desirable that these control points are detected automatically rather than manually. There are many well-developed algorithms that can detect the control points. The Moravec corner detection algorithm is a simple and fast point feature algorithm (Moravec 1977, 1979); however, it is sensitive to rotation such that the points extracted from one image are different from those extracted from the same image that has been rotated (anisotropic response) (Parks and Gravel 2005). The Moravec operator is also highly sensitive to edges, which means that anything that looks like an edge (i.e. noise) may cause the intensity variation to become significant. The intensity variation is the main criterion that detects a corner in this method (Parks and Gravel 2005).

The Harris (Plessey) operator (Harris and Stephens 1988) is another algorithm, which has a higher detection rate than the Moravec operator. The former operator is more robust than the latter one in terms of repeatability (Parks and Gravel 2005). On the other hand, the Harris detector is computationally more costly than the Moravec detector. Since the Harris technique is based on gradient variations, it is also sensitive to noise. Recent modifications have made this method capable of responding isotropically (Parks and

Gravel 2005). Figure 1 shows 563 detected features in a real 3D truss model using the Harris operator.

Recently, the scale-invariant feature transform (SIFT) (Lowe 2004) has become a popular choice for feature detection. SIFT features are invariant to changes in scale and rotation, and partially invariant to changes in 3D viewpoint and illumination. The SIFT operator is also highly discriminative and robust to significant amounts of image noise. Features are identified by finding local extrema in a scale–space representation of the image. For each extremum point, SIFT then computes a gradient orientation histogram over a region around the point producing a 128-element descriptor vector. Although SIFT offers better repeatability than the Moravec or Harris operators, the algorithm is computationally more expensive. Figure 2 shows 600 detected features of the same truss system shown in Figure 1 using the SIFT detector.

3.2. *Feature matching*

After extracting appropriate features, it is time to find the correct matching features in the reference image and the image to be registered. Feature matching can be carried out manually by a human operator; however, it is more desirable to match the extracted features automatically. Automatic feature matching is a critical stage in image registration. Matching inappropriate features will lead to a registration failure.

An initial estimate is necessary to identify the correspondences between the extracted features in the target image and their correct matches in the reference image. The common technique used for making the estimation is calculating a quantitative descriptor for each feature. Next, a distance matrix A is computed, where each A_{ij} element indicates the closeness of the i^{th} feature descriptor in the target image and the j^{th} feature descriptor in the reference image (Ringer and Morris 2001). The initial matching features are selected as the smallest elements of A. There should not be any row or column selected more than once. Ringer and Morris (2001) and Scott and Longuet–Higgis (1991) have provided further details about different descriptor distance matrix calculation methods.

Figures 3 and 4 show examples of 75 putative matched features in two different images of a single truss system using the Harris and SIFT detectors respectively. There are false matches made in these images. In order to improve the correspondence estimation, outliers (defined as incorrect matching features) are supposed to be identified. A very useful technique in this regard is random sample consensus (RANSAC) (Fischler and Bolles 1981).

When there are two images of a single scene taken from different view points, any point in one image lies along a line in the other image (Faugeras 1993). This condition could be imposed as the following constraint:

$$u^T F v = 0, \tag{7}$$

where u and v are the points in the two images, and are expressed in the form of $[x, y, 1]^T$ and T is the transpose operator. The fundamental matrix F uses Equation (7) to relate any two corresponding points of two images that represent the same point of a scene. More details and estimation techniques of this matrix are provided by Torr and Murray (1997), Zhang (1998) and Hartley and Zisserman (2000).

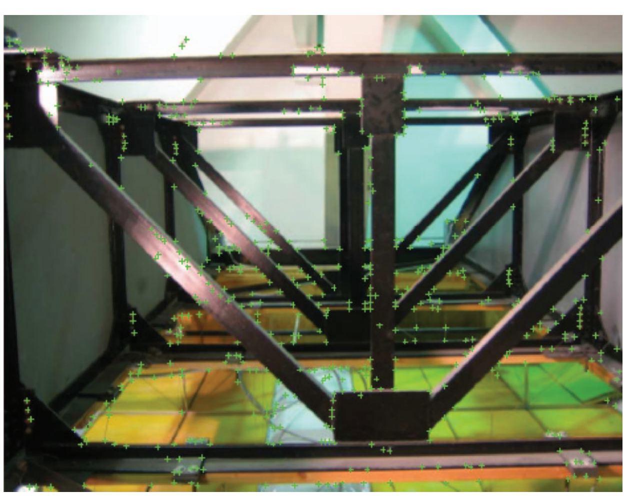


Figure 1. Harris detector – feature points (indicated by the ‘+’ symbol).

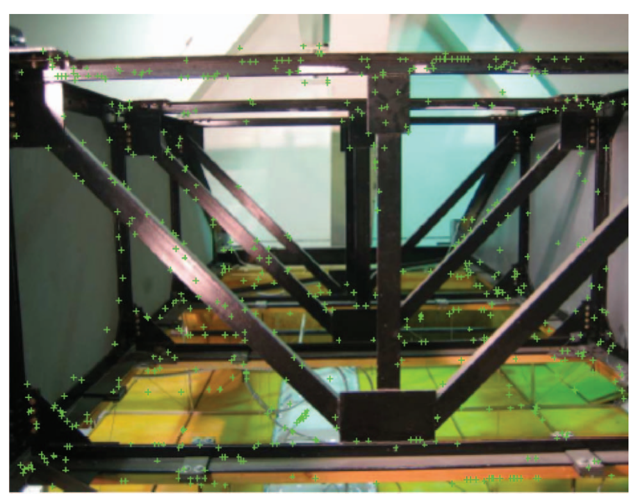


Figure 2. SIFT detector – feature points (indicated by the ‘+’ symbol).

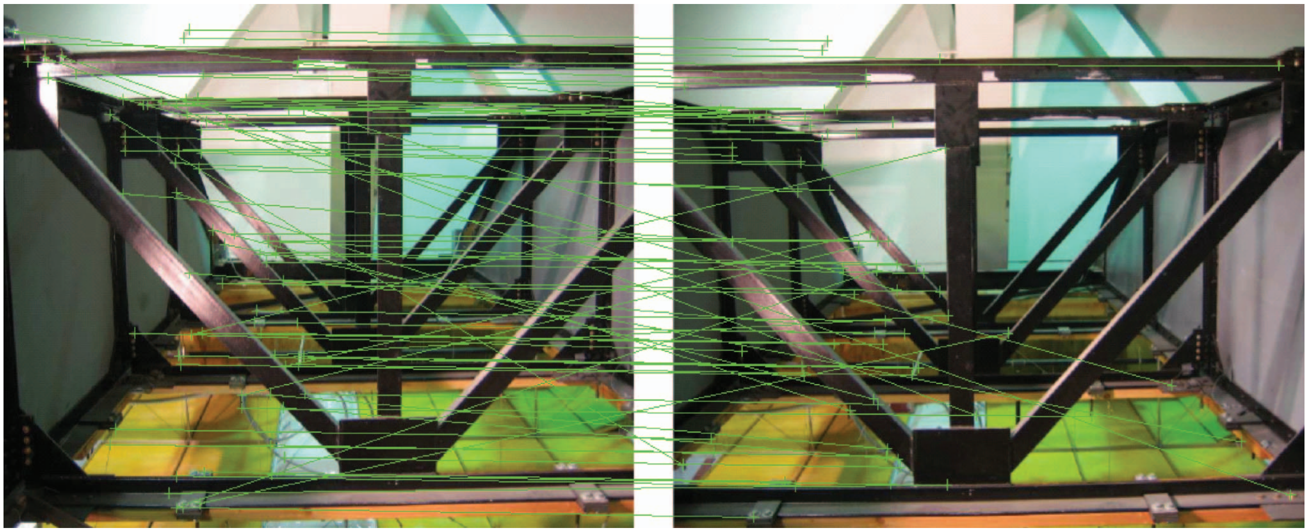


Figure 3. Harris detector – 75 putative matched features in two different images of a single truss system (matched features are connected by matching lines).

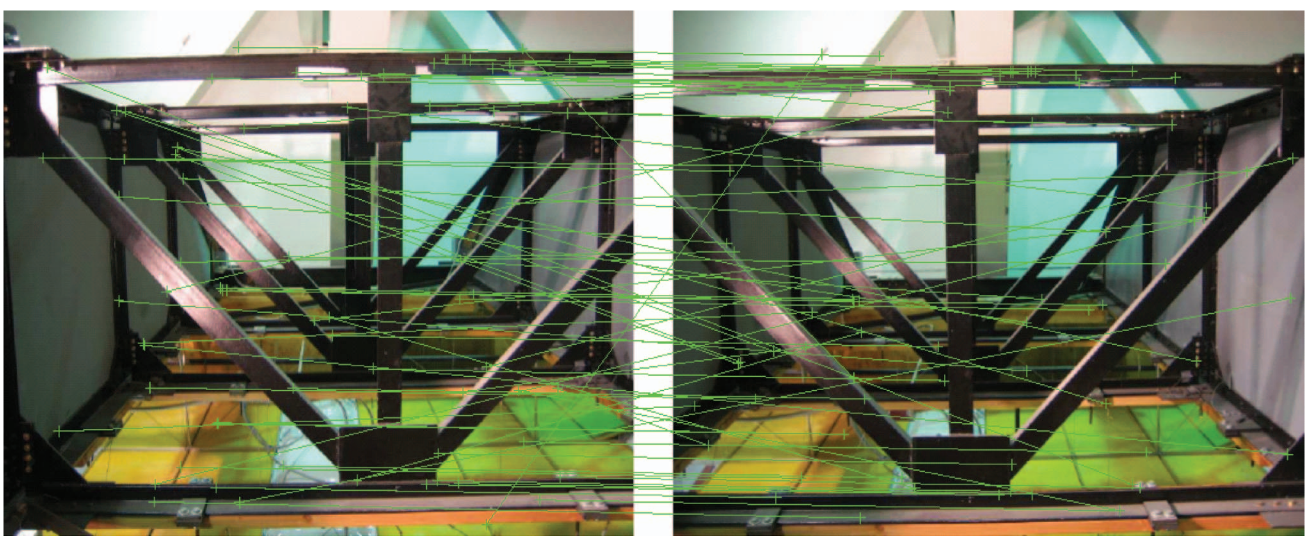


Figure 4. SIFT detector – 75 putative matched features in two different images of a single truss system (matched features are connected by matching lines).

In the RANSAC algorithm, a number of features are randomly chosen to calculate the matrix F. These features are initially estimated to be appropriate matches using the descriptor distance matrix. For each pair of corresponding features, the error is defined as the result of calculating the left-hand side of Equation (7) for the estimated F. Those pairs with errors greater than a threshold are detected as outliers. This procedure is repeated several times until the least amount of total error is calculated, and the minimum number of outliers are detected. Figures 5 and 6 show 75 matched features detected by the Harris and SIFT detectors using the RANSAC algorithm. These two figures contain fewer mismatched features compared with Figures 3 and 4 where the RANSAC algorithm is not used.

This technique can be improved by minimising the median error instead of the total error, also known as the least median squares (LMedS) (Ringer and Morris 2001, Torr and Murray 1997); however, the latter technique is not efficient in the presence of Gaussian noise (Rousseeuw 1987).

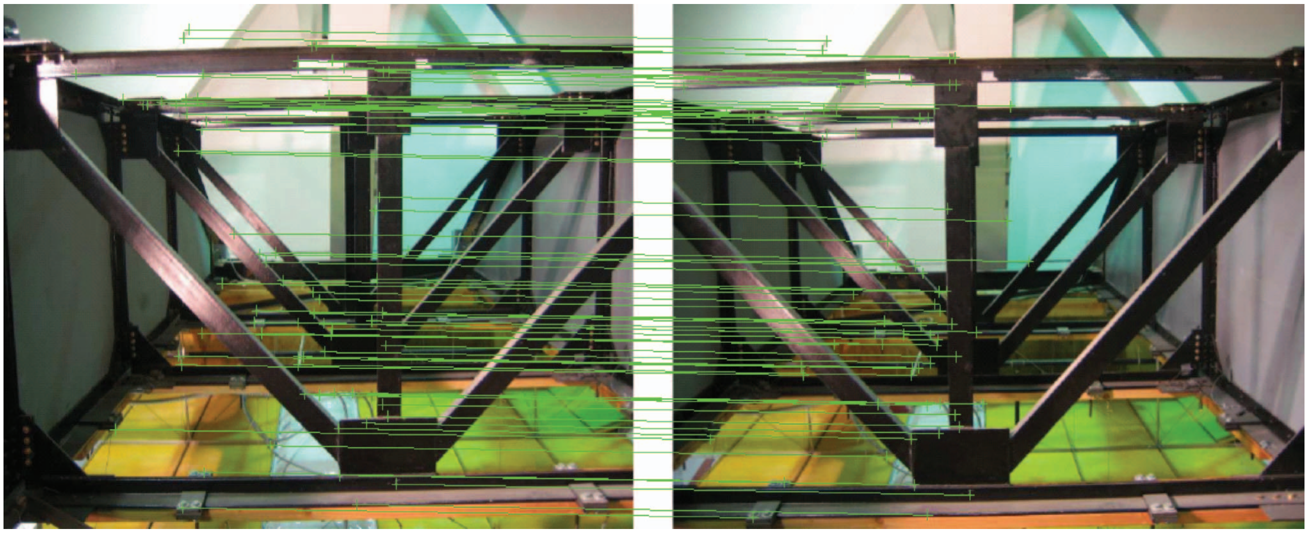


Figure 5. Harris detector – matches consistent with the fundamental matrix.

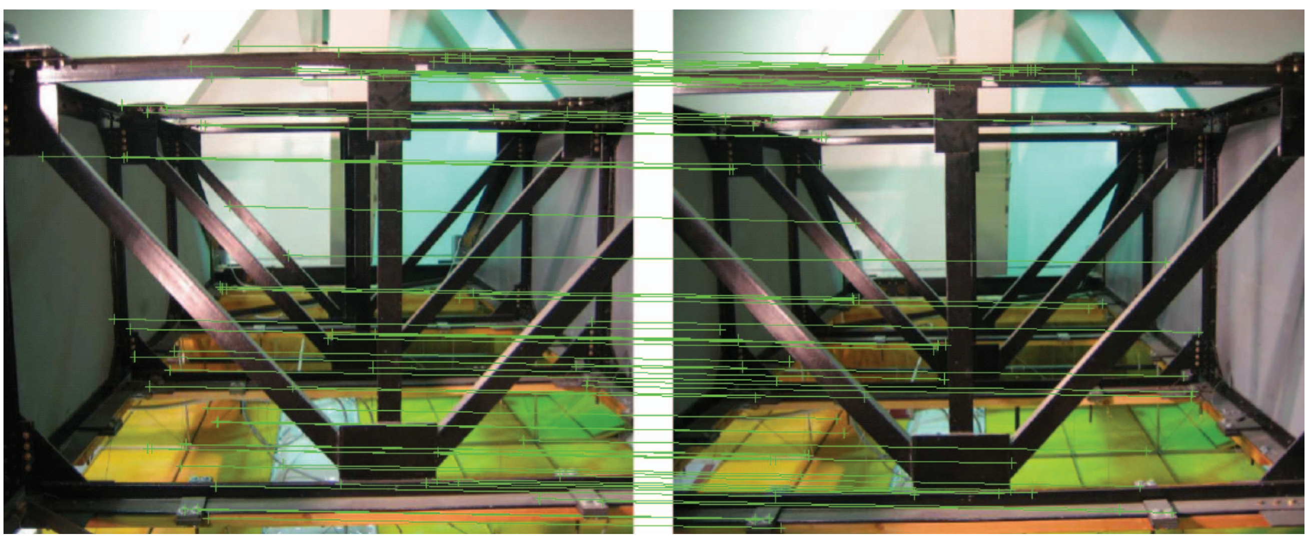


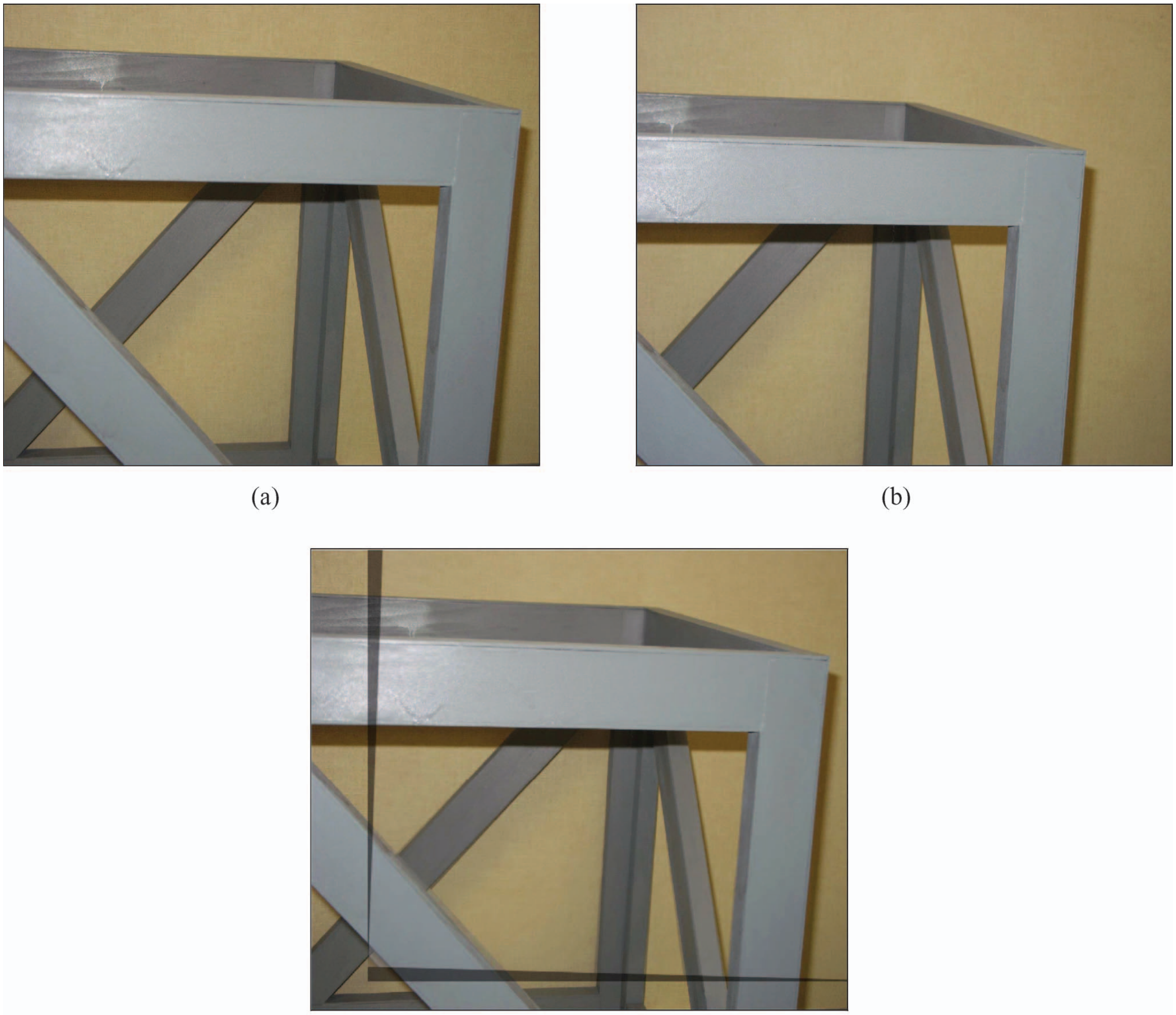
Figure 6. SIFT detector – matches consistent with the fundamental matrix.

3.3. *Transformation and image resampling*

Provided that well-matched features are detected in two images, it is easy to estimate the transformation matrix, which transforms any pixel in the target image into the corresponding pixel in the reference image. Since the pixel coordinates have to be integer numbers, appropriate interpolation techniques are required to calculate the values of the transformed pixels with non-integer coordinates. Figure 7 shows a registration example, where Figures 7(a) and 7(b) show two images taken of a 3D truss model from different orientations, while Figure 7(c) represents the registration of Figure 7(b) onto Figure 7(a). It is worth mentioning that there are generally two types of image registration algorithms, and these are described below.

3.3.1. *Area-based methods*

Area-based methods emphasise feature matching, and are less concerned with feature detection. The structure of the image is analysed through the use of correlation matrices, Fourier properties, etc. (Zitová and Flusser 2003).



(a) (b) (c)

Figure 7. (a) Reference image, (b) image to be registered, and (c) registration of image (b) on image (a).

3.3.2. Feature-based methods

Feature-based methods focus on detecting features (points, lines, corners, line intersections, edges, boundaries, etc.) regardless of image structure (Zitová and Flusser 2003). The feature detection and feature matching algorithms described in §3.1 and §3.2 fall into this type of registration method.

Registration of convex 3D objects has been well-developed in previous studies, but further improvements are needed to register a 3D truss that has occlusions. This is impossible to do with the current registration techniques.

If an image is registered to an existing computer aided design model, it is possible to recover the position of the camera at the time that the image was captured. This positional information can be used to determine how the camera should be repositioned to capture more detailed images of certain areas, etc.

4. Pattern recognition

The aim of pattern recognition is to classify the objects or patterns that have similar attributes into the same class. A scheme of a complete pattern recognition is shown in Figure 8. In this scheme, the first step is data collection. Data sensing can be carried out by a digital camera in this case.

4.1. Segmentation

Segmentation is a set of steps that isolates the patterns that can be potentially classified as the defined defect; however, sometimes it mistakenly picks out patterns that do not belong to the class of potential defects. The aim of segmentation is to reduce the extraneous data about patterns whose classes are not desired to be known. A good segmentation algorithm can help the classifier correctly classify patterns, and it can also affect the type of classifier used. Figure 9(a) shows a corroded column and its background, and Figure 9(b) shows the corroded area segmented from the rest of the

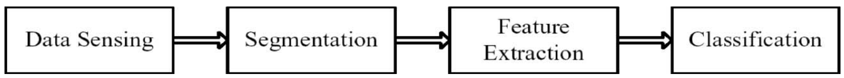
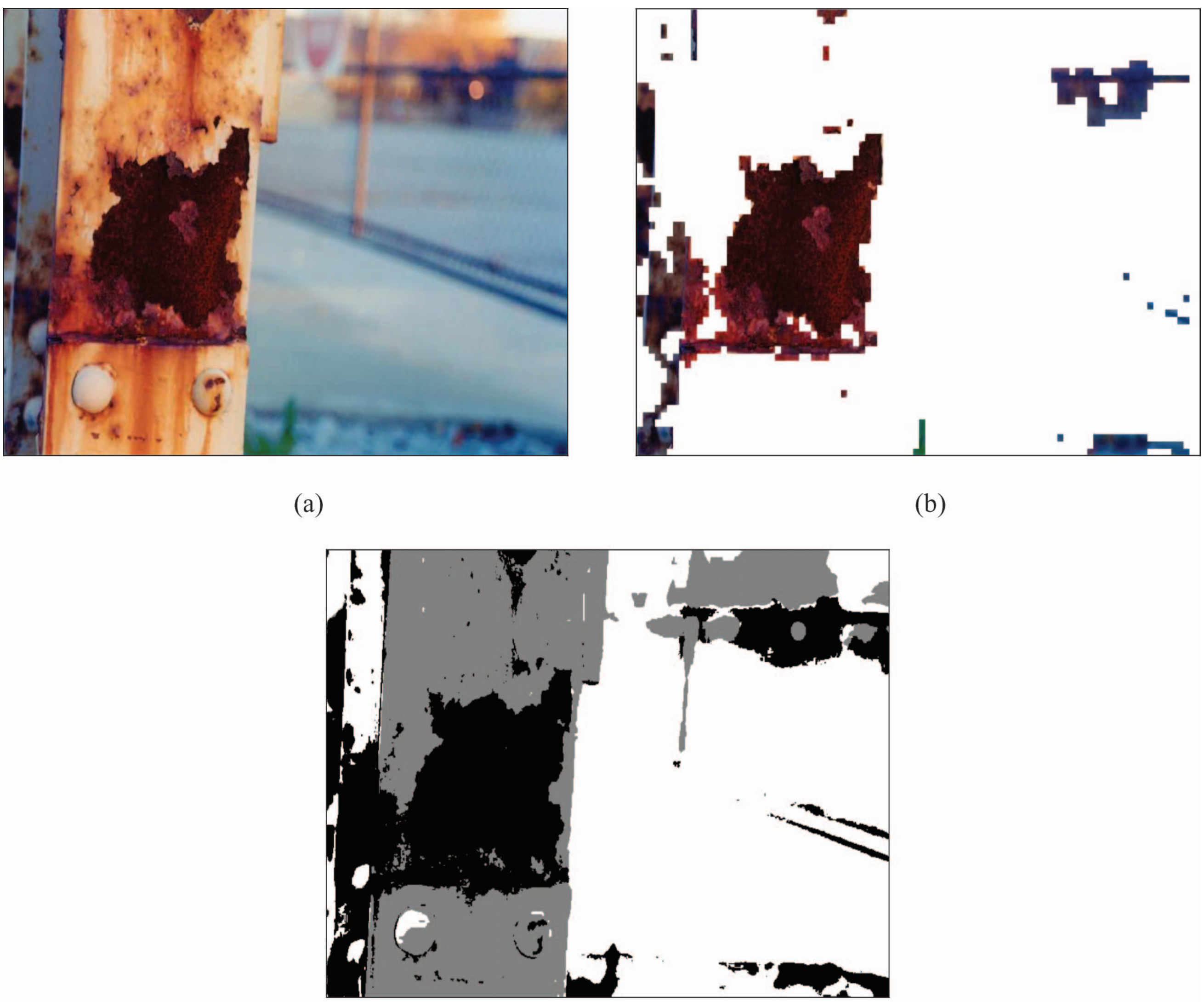


Figure 8. A pattern recognition system scheme.



(a) (b)

(c)

Figure 9. (a) Corroded column and its background, (b) segmentation of the corroded-like area, and (c) classification of pixels using the *k*-means classifier into three classes based on the RGB colour vector of each pixel. (Available in colour online.)

image. Some segmented portions of the image in Figure 9(b) do not belong to the corroded area. Figure 9(c) shows the classification of pixels into three classes using the *k*-means classifier based on the red-green-blue (RGB) colour vector of each pixel. Pixels with the same grey levels (black, white, or grey) belong to the same class.

4.2. *Feature extraction*

After segmenting the patterns of interest, it is time to assign them a set of finite values representing quantitative attributes or properties called features. These features should represent the important characteristics that help identify similar patterns. The process of selecting these suitable attributes is called feature extraction.

According to Hogg (1993), there are five main factors for visual image inspection used by experienced human operators: intensity (a spectral feature), texture (a local spatial feature), size, shape and organisation. In automatic classification of patterns or objects in an image, the spectral and textural attributes are used as features (Sinha *et al*. 2003).

It is possible to assign a feature vector to each pattern in which the elements of the vector are quantitative values representing the extracted features of the pattern. This means that an M-dimensional feature space can be defined where each axis in this space represents a feature and each pattern is one point. It is better if the coordinates are orthogonal in the feature space (it is preferred that the features are independent and also orthogonal). In order to lower the feature vector dimension, it is possible to map the principal features of a pattern from a higher dimensional space to a lower dimensional space by means of a mapping transformation, such as discrete cosine transformation, Fourier transformation or PCA (Karhunen–Loeve transform) (Sinha *et al*. 2003).

PCA is a linear transformation that keeps the subspace with the largest variance, and it needs more computation with respect to the other specified mapping transformations. The goal of this algorithm is to lower the dimensionality of a given data set X from M to L where $L < M$. More information regarding PCA can be found in Jolliffe (2002).

4.3. *Classification*

The last step in a pattern recognition system is decision making or classification. The feature vectors previously extracted for each pattern are inputted into the appropriate classifier, which then outputs the classified patterns. Figure 10 shows the clustering of 7776 pixels of Figure 9(c) plotted in a three-dimensional feature

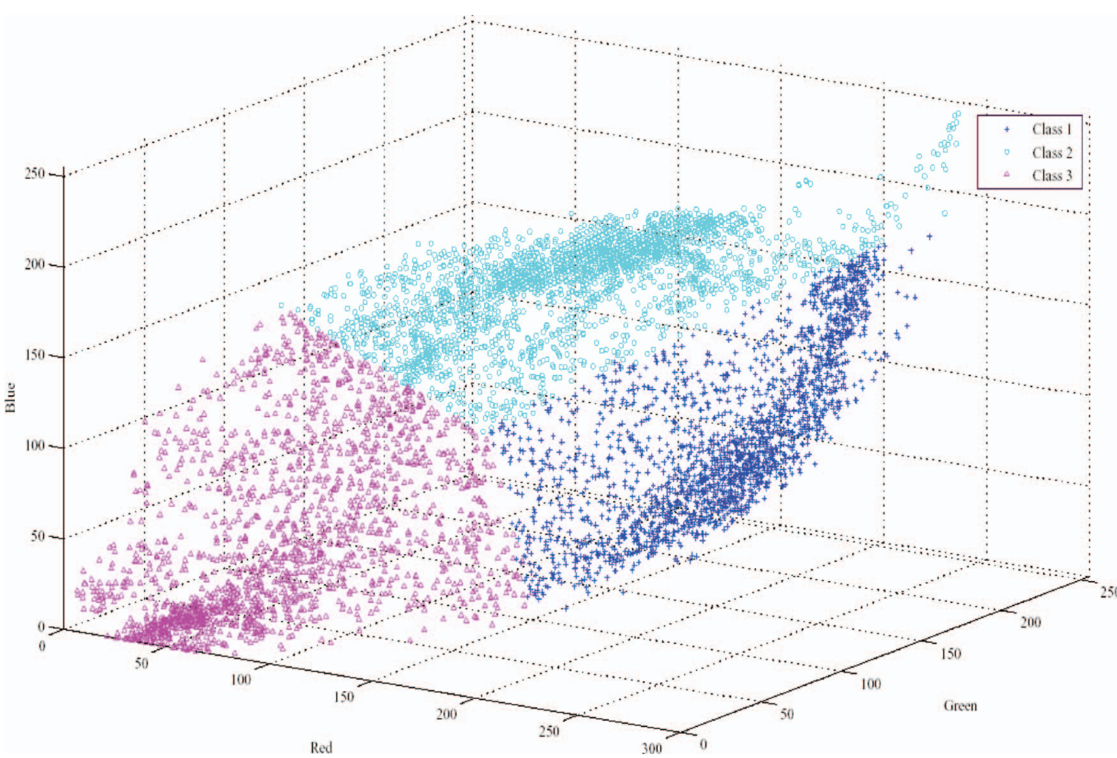


Figure 10. Clustering of pixels of Figure 9(c) in the RGB feature space. (Available in colour online.)

space, the RGB colour space. The two types of classifiers are described below.

4.3.1. Supervised classification

In this type of classification, a set of feature vectors belonging to the known classes is used to train the classifier. The goal of using a training set is to find a relation between the extracted features of the same-class patterns and predict the class of a valid feature vector when its class is unknown. Choosing an appropriate training set is essential to obtaining reasonable and accurate results from supervised classification.

The k-nearest neighbour classifier is an example of supervised classifiers. In this technique, an unclassified pattern is classified, based on the majority of its k nearest neighbours in the feature space. The neighbours are the patterns in the training set. The distance between patterns in the feature space is usually the Euclidian distance. This classifier is very sensitive to noise; however, incrementing k decreases its sensitivity to noise. The appropriate value of k depends on the type of data. If the population of the training set grows enough, the nearest neighbour in the training set represents the class of the unknown pattern. A support vector machine or an artificial neural network can be used as a supervised classification tool in many classification problems. Further details are given by Duda *et al.* (2001).

Neuro-fuzzy systems are promising approaches used by Chae (2001) and Sinha and Fieguth (2006a) to detect defects, including cracks in sewer pipe systems. The performance of the neuro-fuzzy systems proposed in these two researches is better than the regular neural networks and other classical classification algorithms (Chae 2001, Sinha and Fieguth 2006a). Kumar and Taheri (2007) used neuro-fuzzy expert systems in their automated interpretation system for pipeline condition assessment. Neuro-fuzzy systems simultaneously benefit from the data imprecision tolerance (vague definitions) of fuzzy logic systems and the tolerance of the neural networks to noisy data. The easily comprehendible linguistic terms and if-then rules of the fuzzy systems and the learning capabilities of the neural networks are fused into a neuro-fuzzy system (Lee 2005). Different fusions of neural networks and fuzzy systems, which lead to neuro-fuzzy expert systems, are provided by Lee (2005).

4.3.2. Unsupervised classification

There is no training set in an unsupervised classification system. Instead, there is a set of non-classified patterns. The goal is to classify or cluster different patterns of a given data set. This technique is very useful in cases where obtaining an appropriate training set is time consuming or costly. In some cases, a large amount of data can be clustered by an unsupervised classifier, and the class of each cluster can be determined using a supervised classification (Duda and Hart 1973, Duda *et al.* 2001).

A very common unsupervised classifier is the k-means classifier. The goal of this classifier is to cluster patterns into k classes (k is known). In order to achieve this goal, k feature vectors of given patterns are selected randomly as the initial mean of each k class. Each remaining pattern is classified as the class with the nearest mean vector to it. The distance is usually the Euclidean distance. After clustering the data, the mean vector of each class is computed. The patterns are clustered again based on the nearest mean vector. These steps are repeated until the mean vectors do not change or a specific number of iterations is reached. The mean value can be calculated from Equation (8):

$$\mu_i = \frac{1}{N_i} \sum_{j \in C_i} X_j, \tag{8}$$

where C_i is the i^{th} cluster, N_i is the number of members belonging to the i^{th} cluster and X_j is the feature vector of the j^{th} pattern.

The negative aspect of this classifier is the predefined value of k. For automatic clustering of patterns of an image, it is important to find the optimum value of k for that image. Porter and Canagarajah (1996) proposed a way to automatically detect the true cluster number when the number of the clusters is unknown. Within-cluster distance is defined as the sum of all distances between feature vectors and their corresponding cluster mean vectors. Within-cluster distance can be used as a criterion to select the true clustering number. Within-clustering distance D_k can be defined, as indicated by Equation (9), where $d(\mu_i, X_j)$ represents the distance between the mean vector of the i^{th} cluster and feature vector X_j:

$$D_k = \frac{1}{\sum_{m=1}^{k} N_m} \sum_{i=1}^{k} \sum_{j \in C_i} d(\mu_i, X_j). \tag{9}$$

The maximum value for within-cluster distance occurs at $k = 1$; as k increases, D_k rapidly decreases until it reaches the true cluster number, after which D_k decreases very slightly or converges to a constant value. When the first difference of the within-cluster distances is small enough, the true cluster number is found; however, this method requires a threshold. On the other hand, the rapid decrease of D_k before the true cluster number and its gradual decrease after the true cluster number means that the gradient of the 'within-cluster distance' versus 'cluster number' graph has a significant change at the true cluster number. Based on this, the true cluster number is the one that has the maximum value for the second difference of within-cluster distance (Porter and Canagarajah 1996).

In §6 and §7, applicable segmentation techniques and useful features for detection of cracks and corrosion will be reviewed.

5. Wavelet filter bank

The 2D discrete wavelet transform (DWT) of images (Mallat 1989) is a useful technique in many image processing problems, and there are many papers published on this subject. Wavelet transform provides a remarkable understanding of the spatial and frequency characteristics of an image (Gonzalez *et al.* 2004). Since the low frequencies dominate most images, the ability of wavelet transform to repetitively decompose in low frequencies makes it popular for many image analysis tasks (Porter and Canagarajah 1996). In this section, the decomposition and reconstruction of images using wavelet transforms are introduced.

Figure 11 represents a schematic decomposition procedure of an image by 2D DWT. The input to this system is the initial image, the i^{th} approximation; h_φ and h_ψ are low-pass and high-pass decomposition filters, respectively. The words 'Columns' and 'Rows' underneath these filters indicate whether the columns or rows of the input should be convolved with the decomposition filter. Since one-step decomposition of the input with a low-pass and a high-pass filter yields almost a doubled amount of data, a down-sampling (indicated by $2\downarrow$) keeps the amount of data almost the same size as the input. The words 'Columns' and 'Rows' beneath the down-sampling boxes shows that the down-sampling should take place either over columns or rows (which could be done simply by keeping the even-indexed columns or rows).

The result is an $(i+1)^{th}$ approximation, which includes the low frequency characteristics of the input, and it is the most similar output to the input image. There will be three horizontal, vertical, and diagonal details that include the details of the input in the specified directions. These outputs are the wavelet transform coefficients. Since the $(i+1)^{th}$ approximation has the most characteristics of the input image, it can be fed to the decomposition system as an input, and decomposition can take place repeatedly (Misiti *et al.* 2006). Figure 12 shows a 2D wavelet tree for a three-stage decomposition of an image.

As the order of decomposition increases, more details will be decomposed from the image. A two-stage decomposition of a truss model is shown in Figure 13. One can see the horizontal, vertical and

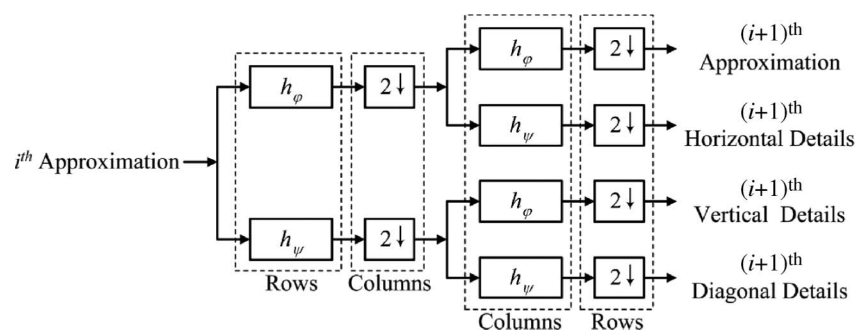


Figure 11. Two-dimensional discrete wavelet transform decomposition scheme.

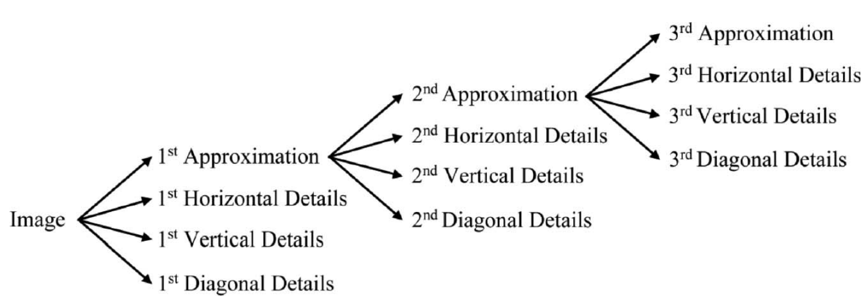


Figure 12. Two-dimensional discrete wavelet transform decomposition scheme. The approximation component of the i^{th} decomposition stage can be decomposed to the $(i+1)^{th}$ approximation, horizontal, vertical and diagonal details.

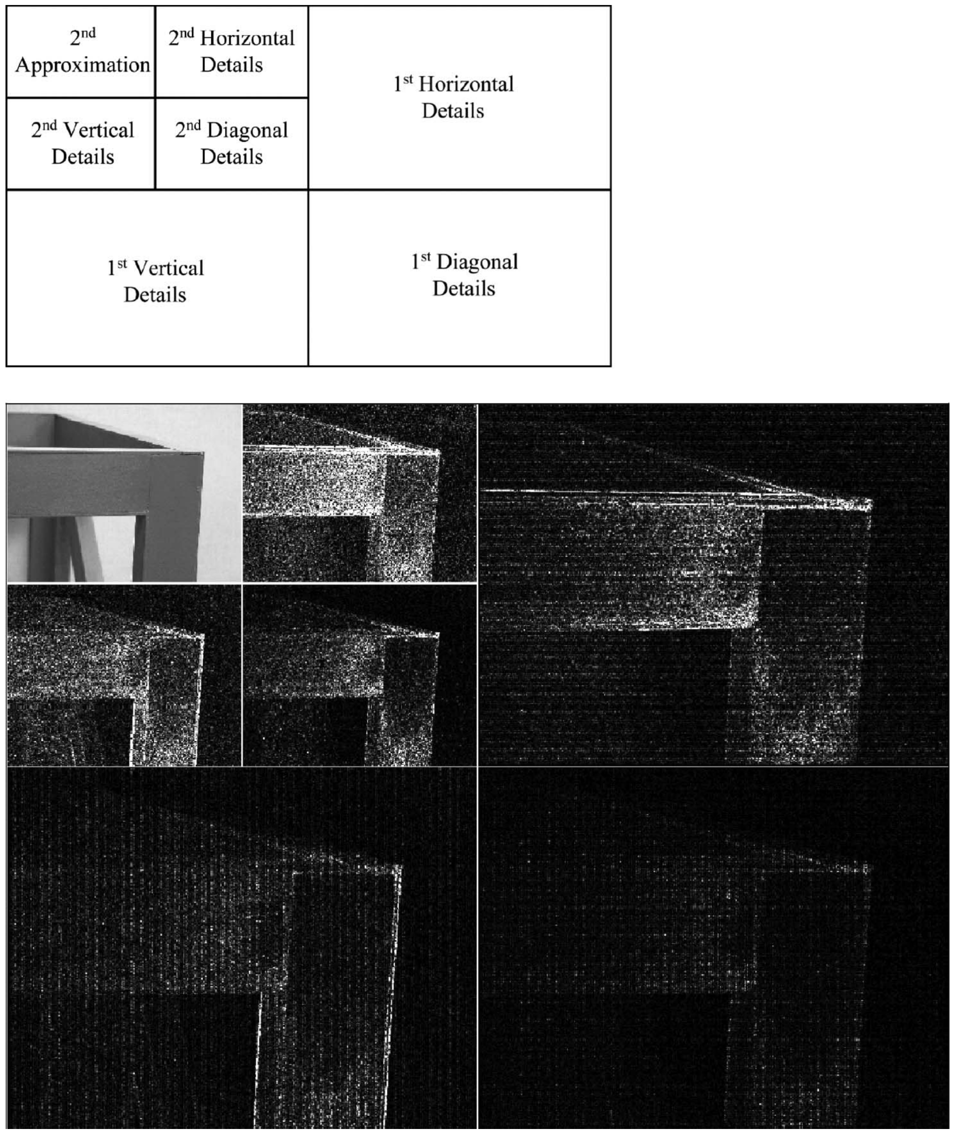


Figure 13. Two-stage DWT decomposition of a truss image.

diagonal details in the two-level decomposition of a truss model in this figure, while the second approximation contains most of the information about the original image. These approximation coefficients and detail coefficients can be used as features for textural analysis of an image. This will be discussed in §7.

The 2D inverse discrete wavelet transform (IDWT) can be used to reconstruct the initial image from the approximation coefficients and detail coefficients as shown in Figure 14. The notation $2\uparrow$ indicates upsampling over rows or columns, which can be done by inserting zeros at odd-indexed rows or columns. Low-pass and high-pass reconstruction filters are denoted as h'_φ and h'_ψ.

The decomposition and reconstruction filters are derived from the scaling function φ and the mother wavelet ψ of a specific wavelet family. The decomposition filters in Figure 13 are based on a Daubechies (1992) wavelet family of order eight. The coefficients for this wavelet family in the form of column vectors (as shown in Equation (10)) and their transposes can be used for column and row convolutions, respectively:

$$h_\varphi = \begin{bmatrix} 0.23038 \\ 0.71485 \\ 0.63088 \\ -0.02798 \\ -0.18703 \\ 0.03084 \\ 0.03288 \\ -0.01060 \end{bmatrix} \quad \text{and} \quad h_\psi = \begin{bmatrix} -0.01060 \\ -0.03288 \\ 0.03084 \\ 0.18703 \\ -0.02798 \\ -0.63088 \\ 0.71485 \\ -0.23038 \end{bmatrix}. \qquad (10)$$

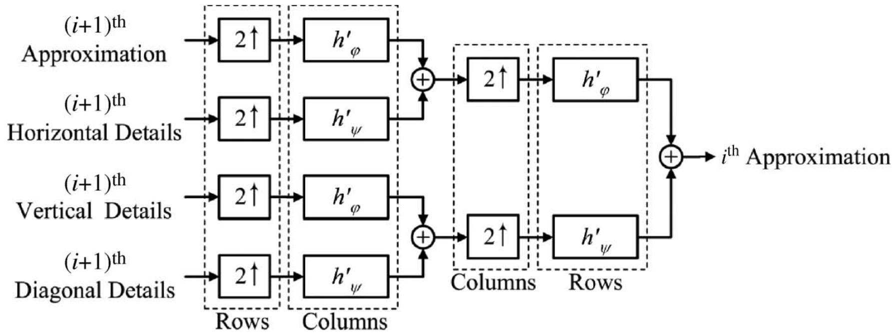


Figure 14. Two-dimensional discrete wavelet transform reconstruction scheme.

Repetitive wavelet decomposition of an image, followed by elimination of its details using a threshold and reconstruction of the edited data, leads to image smoothing, while elimination of the approximations will lead to edge detection. The latter characteristic can be used for crack extraction. Wavelet transform is also used as an image compression tool (DeVore *et al.* 1992, Villasenor *et al.* 1995), and by setting an appropriate threshold its performance is confirmed as a noise removal technique (Chang *et al.* 2000).

6. Crack detection

In this section, two different categories of crack extraction techniques are considered, and the performances of different methods are discussed. The main objective of these techniques is to appropriately segment the regions of interest (crack-like defects) from the rest of the image, followed by feature extraction and decision making to detect the actual cracks. These two categories are based on edge detection and morphological operators. Since edges are often intermixed with cracks, it is hard to classify the cracks from the edges solely based on edge detection methods. Edge detection techniques are more effective when just the defective region is included in the image, or the region of interest is completely segmented from its background. This is not an easy task; however, morphological operations are capable of detecting edges in images that include many non-crack edges. An ideal crack extraction procedure is to derive the initial information by a morphological operation; based on that, the camera can zoom in to capture a closer image of the crack. Finally, the edge detection techniques can be used to extract more information such as the length, the thickness or the orientation of the crack.

6.1. Edge-based techniques

Edge detection techniques can be used to extract crack-like edges in the region of interest where usual edges such as element boundaries do not exist. Two review papers on edge detection techniques are provided by Davis (1975) and Ziou and Tabbone (1998). A comprehensive description of several edge detection techniques is reviewed by Pratt (2001). Any edge detection technique should consist of smoothing, differentiation, and labelling. Smoothing is a preprocessing step that reduces noise and may cause the loss of some edges. An edge can be defined as a discontinuity or sudden change in the image intensity. This is identical to the derivative of an image having local maximum values at the edges. For this purpose, the gradient of an image is an appropriate tool for identifying edges. The gradient vector of a given image $f(x, y)$ is defined as in Equation (11):

$$\nabla \mathbf{f} = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f}{\partial x} \\ \dfrac{\partial f}{\partial y} \end{bmatrix}. \tag{11}$$

The magnitude of the gradient, located at (i, j), can be calculated as indicated by Equation (12):

$$|\nabla \mathbf{f}(i,j)| = (G_x^2(i,j) + G_y^2(i,j))^{\frac{1}{2}}. \tag{12}$$

For computational simplicity, one can approximate the magnitude of the gradient using Equations (13) or (14):

$$|\nabla \mathbf{f}(i,j)| = G_x^2(i,j) + G_y^2(i,j), \tag{13}$$

and

$$|\nabla \mathbf{f}(i,j)| = |G_x(i,j)| + |G_y(i,j)|. \tag{14}$$

The gradient magnitude is zero in areas of constant intensity, whereas in the presence of edges the magnitude is the local maximum. Gradient edge detection can be used to compute the direction of changes as defined in Equation (15):

$$\theta(i,j) = \tan^{-1}\left(\frac{G_y(i,j)}{G_x(i,j)}\right). \tag{15}$$

Common convolution masks (kernels) for digital estimation of G_x and G_y are Sobel (Duda and Hart 1973, Gonzalez *et al.* 2004), Roberts (1965), Prewitt (1970), Frei and Chen (1977) and Canny (1986) edge detection operators. Convolving the initial image with one of the first-order derivative edge detection masks, both vertically and horizontally, generates the approximate gradient magnitude of the pixels. The pixels with values greater than a specified threshold are determined to be edges (labelling). Lower threshold values will lead to detection of more edges, while higher values will cause some edges to be undetected. Different techniques have been proposed to select the appropriate threshold (Abdou 1973, Abdou and Pratt 1979, Henstock and Chelberg 1996). Gonzalez and Woods (1992) proposed an automatic way to compute the global threshold by selecting an initial random threshold T on the histogram of the image. Then μ_1 and μ_2 are computed as the average intensity values of the pixels with intensity values that are greater or less than T, respectively. A new T is then computed as the average of μ_1 and μ_2. This iteration continues until a constant T is achieved.

Another category of the first-order derivative edge detection techniques is based on computing the gradient in more than two orthogonal directions by convolving the initial image with several gradient impulse response arrays and then selecting the maximum value of the convolved images with different templates as shown in Equation (16):

$$G(i,j) = \max\{|G_1(i,j)|, |G_2(i,j)|, \ldots, |G_N(i,j)|\}, \tag{16}$$

where $G_k(i, j)$ is the result of convolving the initial image $f(x, y)$ with the k^{th} gradient response array.

Since the intensity of pixels in an image changes rapidly at the edges (the first derivative has a local maximum), the second derivative will have a zero crossing. The second-order derivative of $f(x, y)$ can be computed by the Laplacian operator as defined in Equation (17):

$$\nabla^2 f(x,y) = \frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2}. \tag{17}$$

In order to compute the second derivative of an image, a window mask is convolved with the image:

$$L(x,y) = f(x,y) * H(x,y). \tag{18}$$

A simple four-neighbour Laplacian mask is:

$$H = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}. \tag{19}$$

The Laplacian is rarely used for edge detection alone because it is very sensitive to noise and cannot detect the direction of edges. Laplacian convolution operators will lead to double-edge detection, which is inappropriate for direct edge detection; however, they can be a complement for other edge detection techniques. Applying a Gaussian smoothing filter to an image and then using the Laplacian of the new image for edge detection yields to Laplacian of the Gaussian operator. Because of its linearity, this detector can be directly applied as the convolution of the initial image with the Laplacian of the Gaussian function:

$$\nabla^2 h(x,y) = -\left(\frac{(x^2+y^2)-\sigma^2}{\sigma^4}\right)\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right). \tag{20}$$

Increasing the window size of the edge detection operator decreases its sensitivity towards noise. Since the Roberts window is the smallest in size, it is very noise-sensitive, and many spots are detected as edges by this operator. The Prewitt operator is weak in detecting diagonal edges (Pal and Pal 1993). The Sobel operator does not have noise sensitivity as it gives more weight to the pixels closer to the pixel of interest, which is located in the middle of the convolution window. Among the operators mentioned above, the Canny operator has the best performance. This technique considers the edges as the local maxima of the derivative of a Gaussian filter. In other words, the smoothing step is imbedded within the operator. Subsequently, the weak edges and the strong edges are extracted by setting two different thresholds. Finally, the strong edges and the weak edges that are connected to strong edges are detected as the real edges. Consequently, less weak edges are falsely detected.

Based on experiments on bridge pavements, Abdel-Qader *et al.* (2003) have concluded that the Canny edge detection technique is more successful in detecting cracks than Sobel and FFT techniques. This result (that the FFT approach has the worst performance) is also confirmed by the authors of the current paper. The FFT approach includes the frequency properties of the image in the frequency domain. The mathematical FFT formulation is shown in Equation (21):

$$F(u,v) = \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\exp\left(-2\pi j\left(\frac{xu}{M}+\frac{yv}{N}\right)\right), \tag{21}$$

where $f(x, y)$ is the $M \times N$ image, x and y are the spatial coordinates and u and v are the transformation coordinates in the frequency domain. Since the FFT is highly sensitive to noise, it is not recommended to be used for the problem in question.

Abdel-Qader *et al.* (2003) have demonstrated that the fast Haar transform performs even better than the Canny detector for detecting cracks in concrete bridge pavements. Haar is the simplest wavelet whose mother wavelet ψ and scaling function φ are shown below:

$$\psi(t) = \begin{cases} 1 & 0 \le t < \frac{1}{2} \\ -1 & \frac{1}{2} \le t < 1 \\ 0 & \text{elsewhere} \end{cases} \tag{22}$$

and

$$\varphi(t) = \begin{cases} 1 & 0 \le t < 1 \\ 0 & \text{elsewhere} \end{cases}. \tag{23}$$

The decomposition and reconstruction filters for this wavelet family are:

$$h_{\varphi} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad h_{\psi} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}. \tag{24}$$

Abdel-Qader *et al.* (2003) used the Haar wavelet to get the one-level decomposition of an image, as described in §5. Then, the three details are combined to generate the magnitude image. The threshold is defined as the average intensity value of all pixels in the captured images. The overall accuracy of this technique is 86%, as reported by Abdel-Qader *et al.* (2003). Because the effect of light and the contrast of each image are not considered independently when choosing the threshold, this thresholding is inappropriate. By selecting an independent threshold for each image, a better detection rate is expected. In this approach, no other classification is used to detect cracks from non-crack edges.

Mallat and Zhong (1992) have demonstrated that the local maxima of an image wavelet transform can be used to extract and analyse multi-scale edges. Siegel and Gunatilake (1998) have used the wavelet filter bank for detecting cracks in aircraft surfaces, where they used a cubic spline and its first derivative as scaling and wavelet functions. This wavelet transform is equivalent to applying a smoothing filter on the image followed by taking the derivative of the smoothed image, which is identical to a classical edge detection procedure.

Siegel and Gunatilake (1998) applied a three-level decomposition on the region of interest; however, the decomposition algorithm is slightly different from what is described in §5. This decomposition is defined as applying the high-pass decomposition filter, g_i (mother wavelet function), once to the rows and once to the columns separately, which leads to W_{yi} and W_{xi}, respectively, where i is the decomposition level. The low-pass decomposition filter, h_i (scaling function), is applied to the rows and the columns. The wavelet and scaling decomposition filters for the specified wavelet at the three levels are shown below:

$$g_1 = [2, -2], \tag{25}$$

$$g_2 = [2, 0, -2], \tag{26}$$

$$g_3 = [2, 0, 0, 0, -2], \tag{27}$$

$$h_1 = [0.125, 0.375, 0.375, 0.125], \tag{28}$$

$$h_2 = [0.125, 0, 0.375, 0, 0.375, 0, 0.125], \tag{29}$$

and

$$h_3 = [0.125, 0, 0, 0, 0.375, 0, 0, 0, 0, 0.375, 0, 0, 0, 0.125]. \tag{30}$$

The schematic procedure described above is presented in Figure 15. In order to extract crack-like edges, the magnitude image M_i is computed for each level as:

$$M_i = \sqrt{W_{xi}^2 + W_{yi}^2}. \tag{31}$$

By choosing a dynamic threshold based on the histogram of each magnitude image, pixel values above the threshold are detected as edge points. Since the direction of a crack varies smoothly, edge points are linked, based on eight neighbours if the difference of the corresponding angles is less than a specific angle

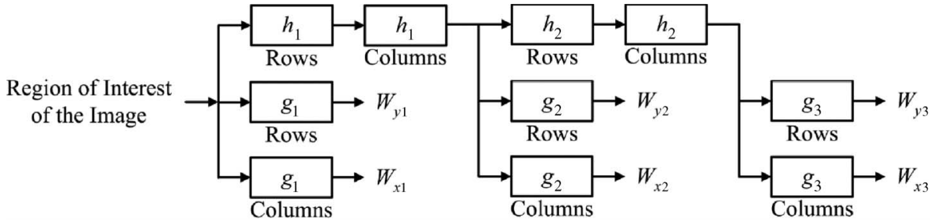


Figure 15. Multi-resolution decomposition of the image used by Siegel and Gunatilake (1998).

(Siegel and Gunatilake 1998). For this purpose, the angle image of each level A_i is defined as:

$$A_i = \arctan\left(\frac{W_{yi}}{W_{xi}}\right). \tag{32}$$

A one hidden-layer neural network consisting of four neural units, six input units and one output unit is used to classify cracks from non-crack edges. On aircraft surfaces, the cracks are smaller than non-crack edges (e.g. scratches). On the other hand, an edge that only appears in the first-decomposition level is smaller than an edge that appears in the first two levels. Similarly, the latter is smaller than an edge that appears in all levels of the decomposition. This important characteristic, which was introduced as 'propagation depth' by Siegel and Gunatilake (1998), is considered to be one of the selected features. The propagation depth represents the number of decomposition levels in which a specific edge has appeared, and also conveys the size information of the edge. A propagation depth is assigned to each edge that appears in the first decomposition level. For this reason, a 'coarse-to-fine edge linking process' is used, which provides information about an edge from a coarse resolution to a fine resolution. The features assigned to each edge in the first decomposition level, as defined by Siegel and Gunatilake (1998), are:

(1) Computed propagation depth number;
(2) Number of pixels constituting the edge;
(3) Average wavelet magnitude of the edge pixels;
(4) Direction of pixels constituting the edge in level one;
(5) The signs of $\sum W_{x1}$ and $\sum W_{y1}$ for all pixels belonging to the edge; and
(6) Average wavelet magnitude of linked edges in levels two and three during coarse-to-fine edge linking processes.

The accuracy of this technique in crack detection is 71.5% (Siegel and Gunatilake 1998). The smaller accuracy of this study with respect to the one carried out by Abdel-Qader *et al*. (2003) is due to the level of complexity of the problem. This technique is highly dependent on the direction of light during the image acquisition. The steps necessary to obtain better performance of the classification process are: selecting additional features, capturing more images of a surface with different camera orientations (in order to gather and enrich the data with different light directions) and also increasing the population of the training set (Siegel and Gunatilake 1998).

None of the techniques discussed above deal with the problem of major non-defect edges such as structural member edges or background crack-like objects. Another set of techniques that can overcome this shortcoming is discussed in the following section.

6.2. Morphological techniques

Morphological image processing extracts useful information about the objects of an image based on mathematical morphology. The foundation of morphological image processing is based on previous studies of Minkowski (1903) and Matheron (1975) on set algebra and topology, respectively (Pratt 2001). Morphological techniques can be applied to binary or grey-scale images. Although morphological operations are also discussed in the context of colour image processing (Comer and Delp 1999, Al-Otum and Munawer 2003, Yu *et al*. 2004), the grey-scale operations that are useful for segmenting cracks from the rest of an image are introduced here. Figure 16(a) shows a vertical crack on a steel strip caused by a tensile rupture, and Figure 16(b) shows a horizontal crack on a rebar caused by a torsional rupture. The results of performing different morphological operations on these two images are presented later in this paper to give the reader a better understanding of the applications of the described operations.

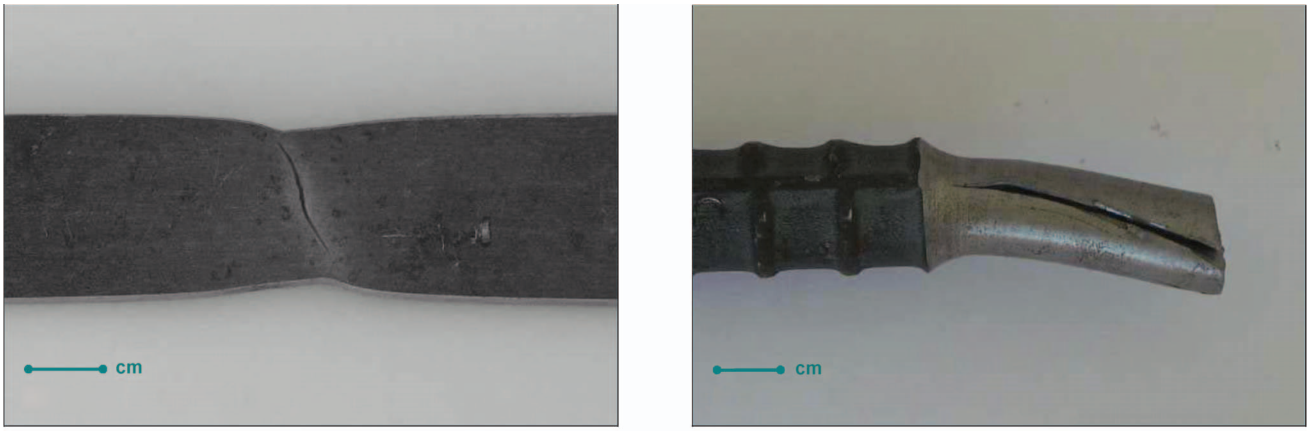
Morphological image processing generally can be used in image filtering, image sharpening or smoothing, noise removal, image segmentation, edge detection, feature detection, defect detection, preprocessing and postprocessing tasks. A brief discussion of some definitions used in morphological approaches follows.

6.2.1. Dilation

The grey-scale dilation of image I and the structuring element S is defined as:

$$\begin{aligned}(I \oplus S)\,(x,y) \;=\; \max\,[I\,(x-x',\; y-y') \\ +\; S(x',\; y') \mid (x',\; y') \in\; Ds],\end{aligned} \tag{33}$$

where D_S, a binary matrix, is the domain of S (the structuring element) and defines which neighbouring pixels are included in the maximum function. In the case of non-flat structuring elements, D_S indicates the pixels included in the maximum function as well as their weights (D_S is not binary in this case). During the dilation process, which is similar to the convolution process, $I\,(x, y)$ is assumed to be $-\infty$ for $(x, y) \notin D_S$. In the case of flat structuring elements, $S\,(x', y') = 0$ for $(x', y') \in D_S$. Flat structuring elements are usually used for grey-scale morphological operations. Visually, dilation expands the bright portions of the image (Salembier 1990). The results of applying this operation to the images in Figure 16 are shown in



(a) (b)

Figure 16. (a) Vertical crack caused by a tensile rupture of a steel strip, and (b) horizontal crack caused by a torsional rupture of a steel rebar.

Figure 18. Figures 17(a) and (b) show the domain of two flat structuring elements used in Figures 18, 19, 20, 21, 22, 24 and 25. Number 1 in Figures 17(a) and (b) shows the pixels to be included in the morphological operation.

6.2.2. *Erosion*

Similar to above, the grey-scale erosion is defined as:

$$(I \ominus S)(x,y) = \min [I(x + x', y + y') - S(x', y') \mid (x', y') \in Ds], \quad (34)$$

where $I(x, y)$ is assumed to be $+\infty$ for $(x, y) \notin D_S$. In the case of flat structuring elements, $S(x', y') = 0$ for $(x', y') \in D_s$. In fact, erosion shrinks the bright portions of a given image (Salembier 1990). The results of applying this operation to the images in Figure 16 are shown in Figure 19.

6.2.3. *Morphological gradient*

The morphological gradient is defined as the dilated image minus the eroded version of the image, and it can be used to detect edges as it represents the local variations of an image (Gonzalez *et al*. 2004). Figure 20 shows the results of applying this operation to the images in Figure 16.

6.2.4. *Opening*

The grey-scale opening of an image I by structuring element S can be written as:

$$I \circ S = (I \ominus S) \oplus S. \quad (35)$$

1	1	1	1	1

(a)

1
1
1
1
1
1
1
1
1

(b)

Figure 17. (a) 1 × 5 flat structuring element domain used in Figures 18–22, 24 and 25, and (b) 9 × 1 flat structuring element domain used in Figures 18–22, 24 and 25.

Figure 21 shows the results of applying the opening operation to the images in Figure 16. Sinha and Fieguth (2006b) detected the defects in underground pipe images by thresholding the morphological opening of the pipe images using different structuring elements.

6.2.5. *Closing*

Similarly, the closing in grey-scale is defined as:

$$I \bullet S = (I \oplus S) \ominus S. \quad (36)$$

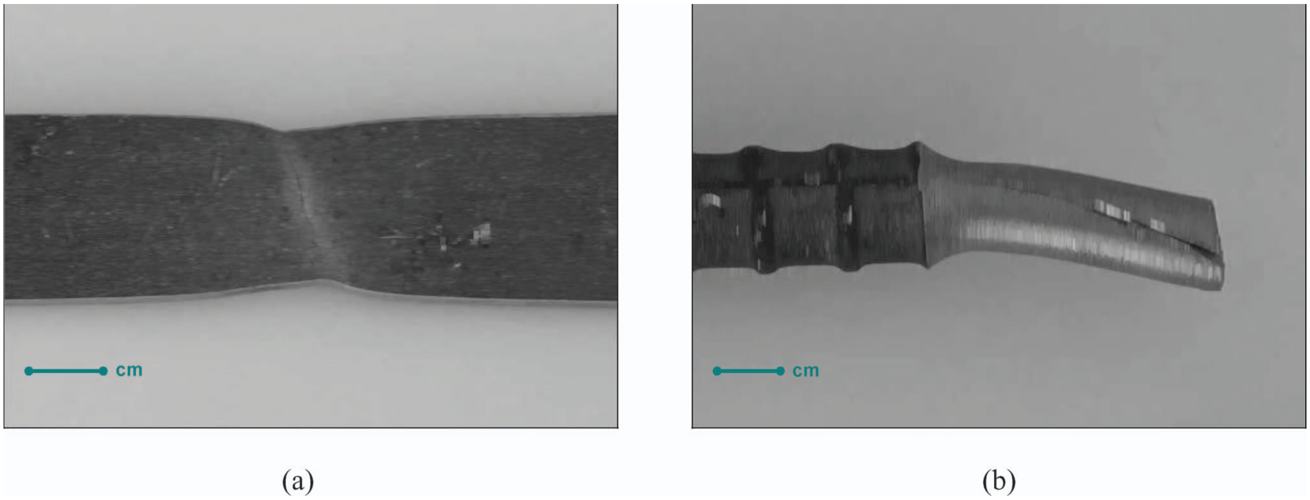


Figure 18. Dilation performed by: (a) a 1 × 5 structuring element, and (b) a 9 × 1 structuring element.

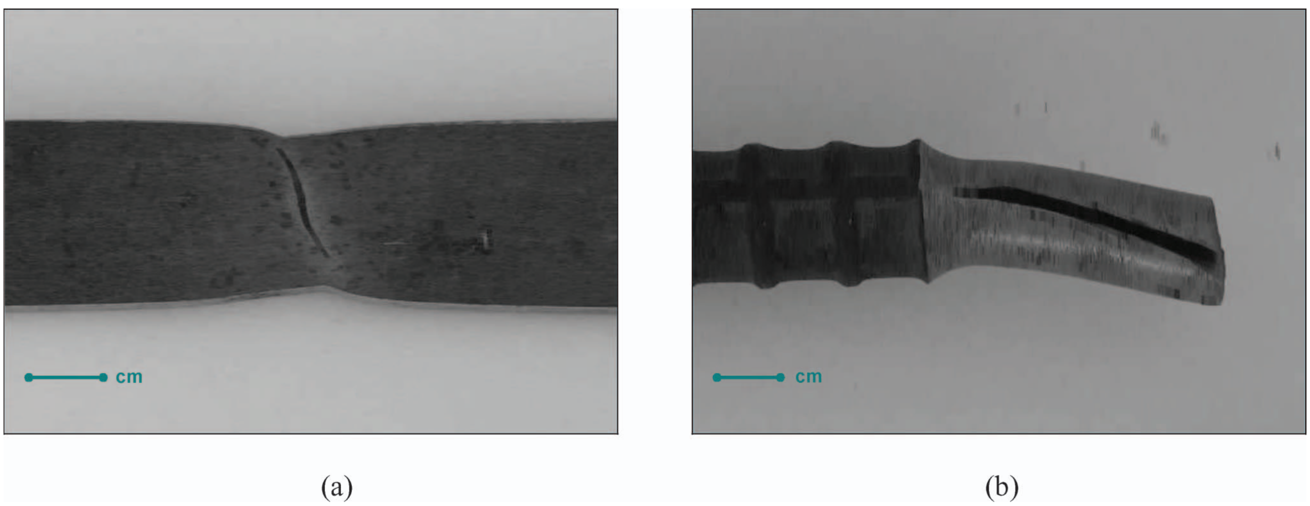


Figure 19. Erosion performed by: (a) a 1 × 5 structuring element, and (b) a 9 × 1 structuring element.

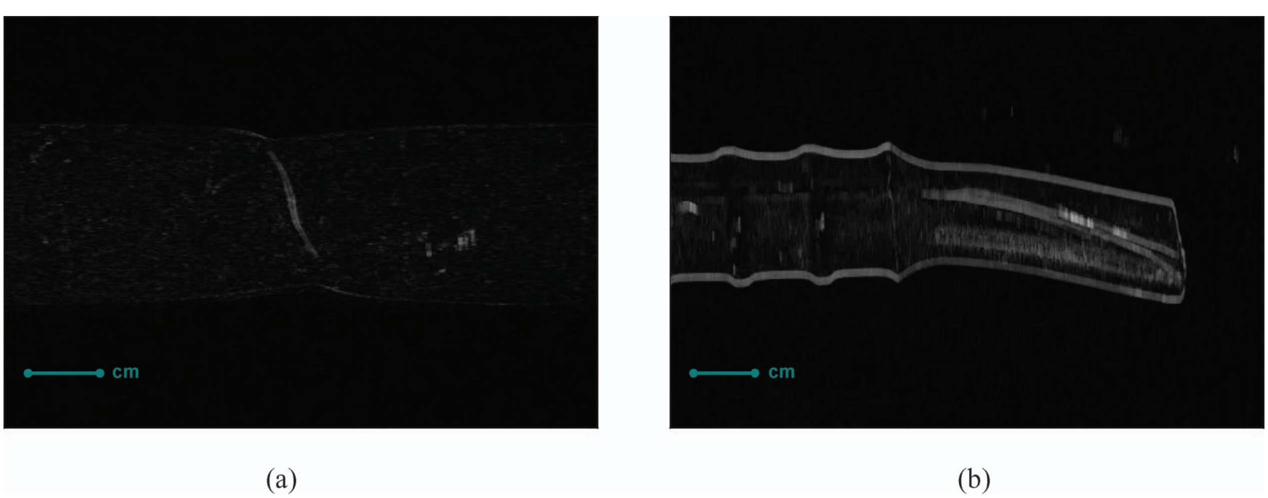


Figure 20. Morphological gradient performed by: (a) a 1 × 5 structuring element, and (b) a 9 × 1 structuring element.

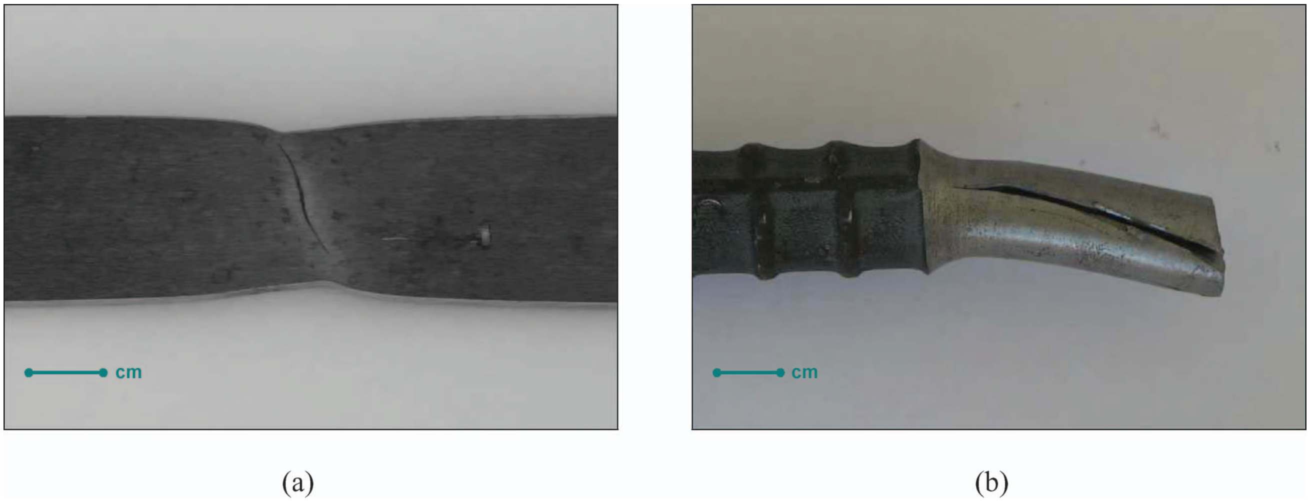


Figure 21. Opening performed by: (a) a 1 × 5 structuring element, and (b) a 9 × 1 structuring element.

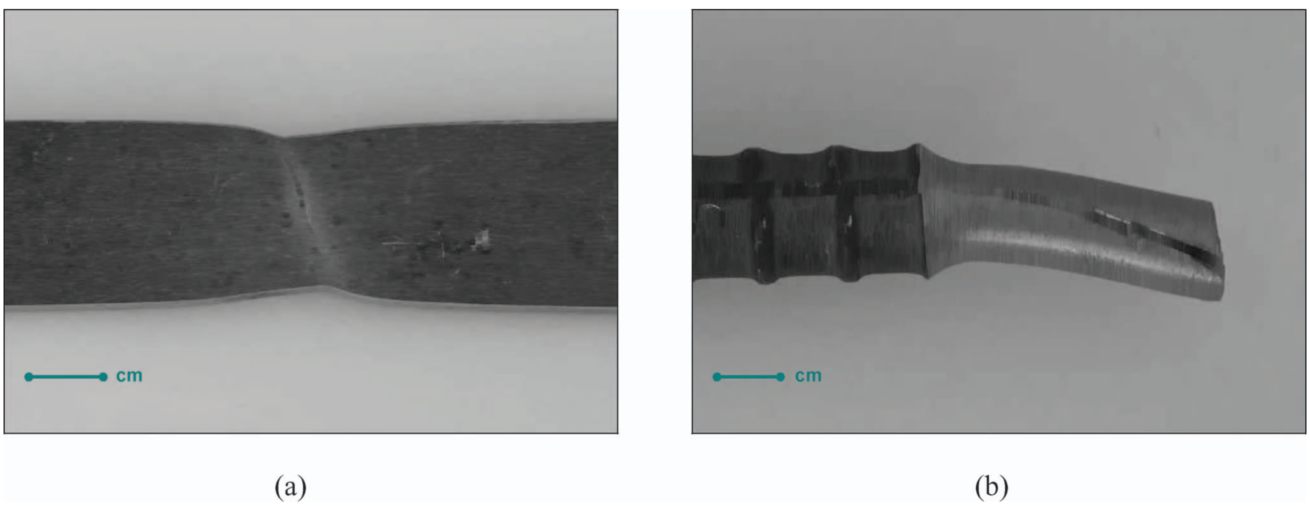


Figure 22. Closing performed by: (a) a 1 × 5 structuring element, and (b) a 9 × 1 structuring element.

Closing is applied to the images in Figure 16, and the results are shown in Figure 22.

While opening is usually used to eliminate sharp bright details, closing is used to remove dark details provided that the structuring element is larger than the details. These properties make the combination of opening and closing very suitable for noise removal and image blurring (Gonzalez *et al.* 2004). Figure 23 shows the one-dimensional opening and closing operations. While the curve represents the grey-scale level, the circles (structuring elements) beneath the curve pushing it up illustrate the opening operation, and the circles above the curve pushing it down represent the closing operation.

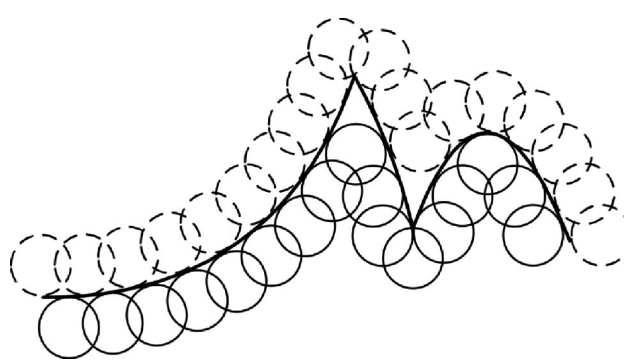


Figure 23. One-dimensional scheme of the opening and closing operations. The curve represents the grey-scale level, the circles (structuring elements) beneath the curve pushing it up illustrate the opening operation, and the circles above the curve pushing it down represent the closing operation.

In Figure 23, higher values of the curve show brighter pixels. One can conclude from Figure 23 that the subtraction of the opened image from its original version will result in the detection of bright defects. This operation is called top-hat transform (Serra 1982, Meyer 1986) and its formula is:

$$T = I - (I \circ S). \tag{37}$$

The dual form of the above equation is called bottom-hat, which is appropriate for detecting dark defects (see Figure 23). Its formulation is:

$$T = (I \bullet S) - I. \tag{38}$$

An example of bottom-hat operation applied to the images in Figure 16 is shown in Figure 24.

These two transformations can also be used for contrast enhancement. Giakoumis *et al*. (2006) detected the cracks in digitised paintings by thresholding the output of the top-hat transform. The shape and size of the structuring element depend on the defect of interest. The structuring element has an important role in defect detection using the above morphological image processing techniques. For example, to detect small defects such as edges, small structuring elements, preferably flat ones, should be used (Salembier 1990).

Salembier (1990) has proposed and compared different algorithms to improve morphological defect detection based on top-hat and bottom-hat transformations. It was concluded that the algorithms shown in Equations (39) and (40) can be used to detect bright and dark defects respectively:

$$T = I - \min\,[(I \bullet S) \circ S, I], \tag{39}$$

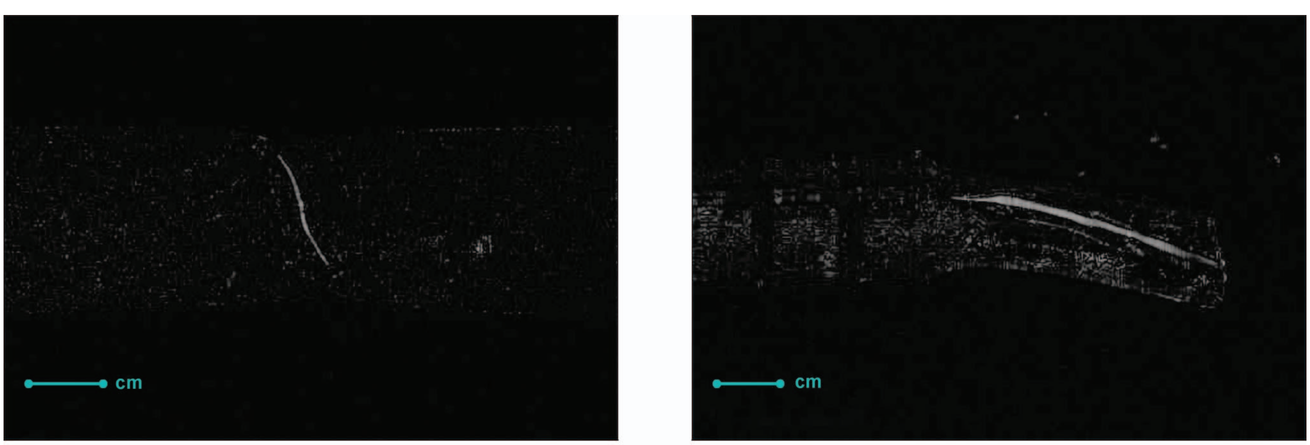
and

$$T = \max\,[(I \circ S) \bullet S, I] - I. \tag{40}$$

Applying Equation (40) to the images in Figure 16 leads to the segmentation of dark cracks, as shown in Figure 25. After postprocessing, the final segmented cracks are shown in Figure 26.

Nieniewski *et al*. (1999) used the above equations to extract cracks in ferrites. They used two flat structuring elements: a 1 × 5 row element for the vertical cracks and a 5 × 1 column element for the horizontal ones. The authors of the current study applied a square structuring element to truss model images to detect cracks; the results after noise removal were excellent. Even in the presence of occlusions and different backgrounds, the algorithm is highly successful provided that the camera captures high resolution images and focuses its lens on the specific structural member of interest. Now-a-days almost all commercial digital cameras have active or passive auto-focusing capabilities that can assist in the acquisition of suitable images (Schlag *et al*. 1983).

Figure 27 represents examples of applying the above technique on actual images. Figures 27(a), (c), (e), (g) and (i) are the original real steel structural members. Figure 27(i) is a magnification of Figure 27(g) to give a better view of the structure's crack. The true cracks that were segmented in Figures 27(b), (d), (f), (h) and (j) are white and the other segmented objects are non-white (after postprocessing and noise removal). The domain of structuring elements that were used to segment the cracks are as follows: a 70 × 70 matrix with (both main and minor) diagonal



(a) (b)

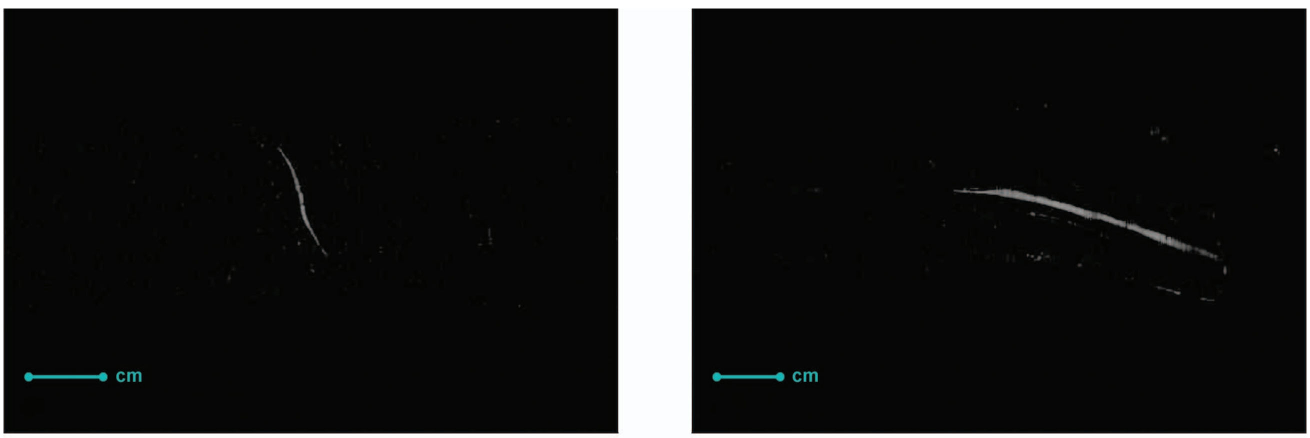
Figure 24. Bottom-hat operation performed by: (a) a 1 × 5 structuring element, and (b) a 9 × 1 structuring element.

members of 1 and non-diagonal members of 0 (Figure 27(b)), a 10 × 10 matrix with ones in the minor diagonal and zeros everywhere else (Figure 27(d)), a unit matrix of 7 × 7 (Figure 27(f)) and a 1 × 5 structuring element (Figures 27(h) and (j)).

The challenge is to find the appropriate size and format of the structuring element. When the structuring element has a line format, it can segment cracks that are perpendicular to it (see the structuring elements used in Figures 26 and 27). A good example of such a study is presented by Sinha and Fieguth (2006b), in which they tried to find the optimal size of structuring elements to segment and classify cracks, holes, laterals and joints in underground pipe images.

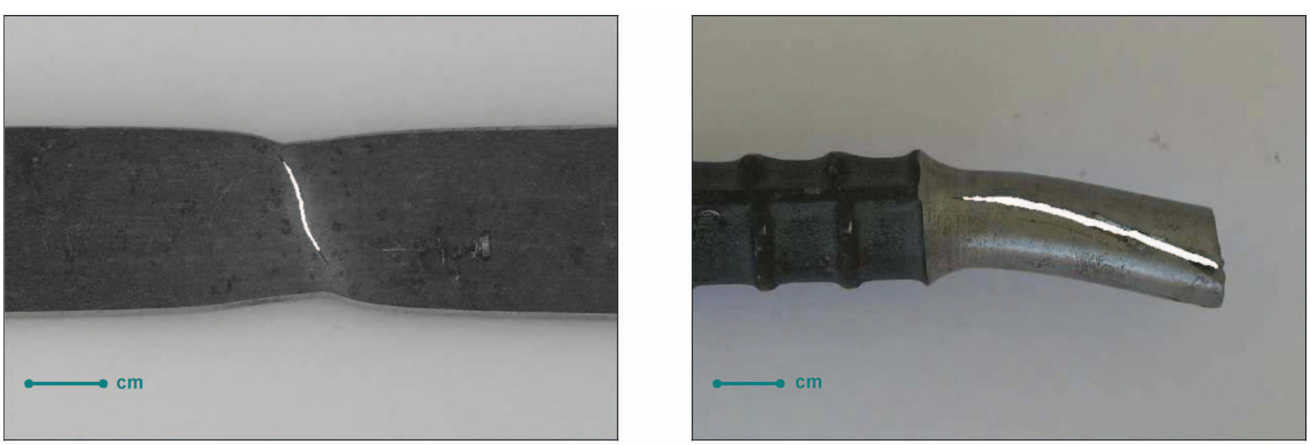
The morphological operations discussed above segment the cracks more efficiently than the edge detection operators reviewed in §6.1. Edge-based techniques extract all the edges in an image, which makes the classification task harder. Basically, edge-based techniques will generate more noise than morphological techniques (compare Figures 25 and 28).

After extracting reliable, independent and discriminating set of features from the segmented objects, it is the classifier's task to label each of the segmented objects as crack or non-crack (§4). Sinha (2000) has used area, number of objects, major axis length, minor axis length, mean and variance of pixels



(a) (b)

Figure 25. (a) Vertical dark crack segmented using morphological techniques, and (b) horizontal dark crack segmented using morphological techniques.



(a) (b)

Figure 26. Crack segmentation. (a) Vertical crack segmentation, and (b) horizontal crack segmentation.

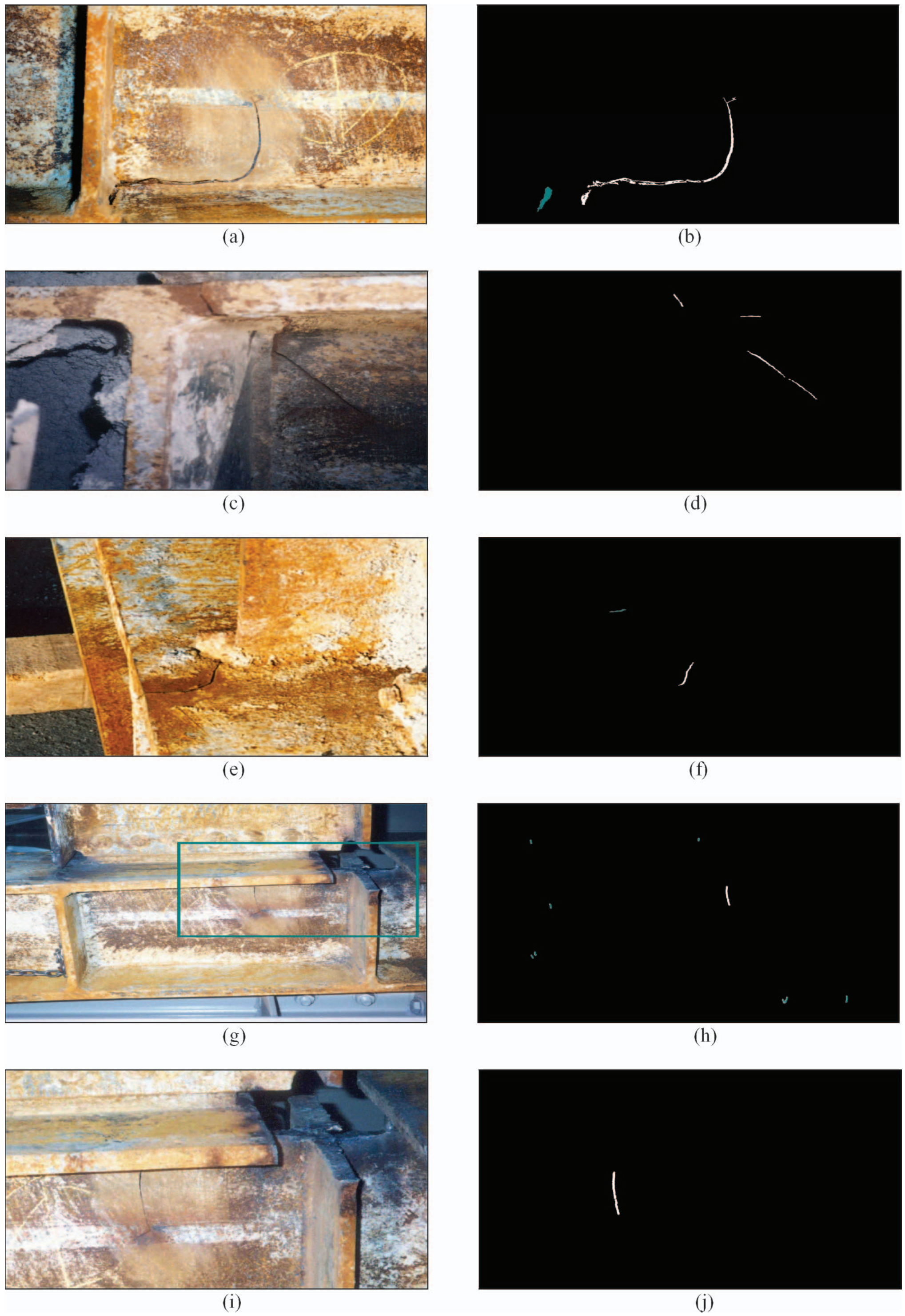


Figure 27. Examples of crack segmentation in real structures: (a) original structural member 1, (b) crack segmentation of image (a) using a 70×70 matrix with (both main and minor) diagonal members of 1 and non-diagonal members of 0 as the structuring element, (c) original structural member 2, (d) crack segmentation of image (c) using a 10×10 matrix with ones on the minor diagonal and zeros everywhere else as the structuring element, (e) original structural member 3, (f) crack segmentation of image (e) using a unit matrix of 7×7 as the structuring element, (g) original structural member 4, (h) crack segmentation of image (g) using a 1×5 structuring element, (i) magnification of image (g), and (j) crack segmentation of image (i) using a 1×5 structuring element.

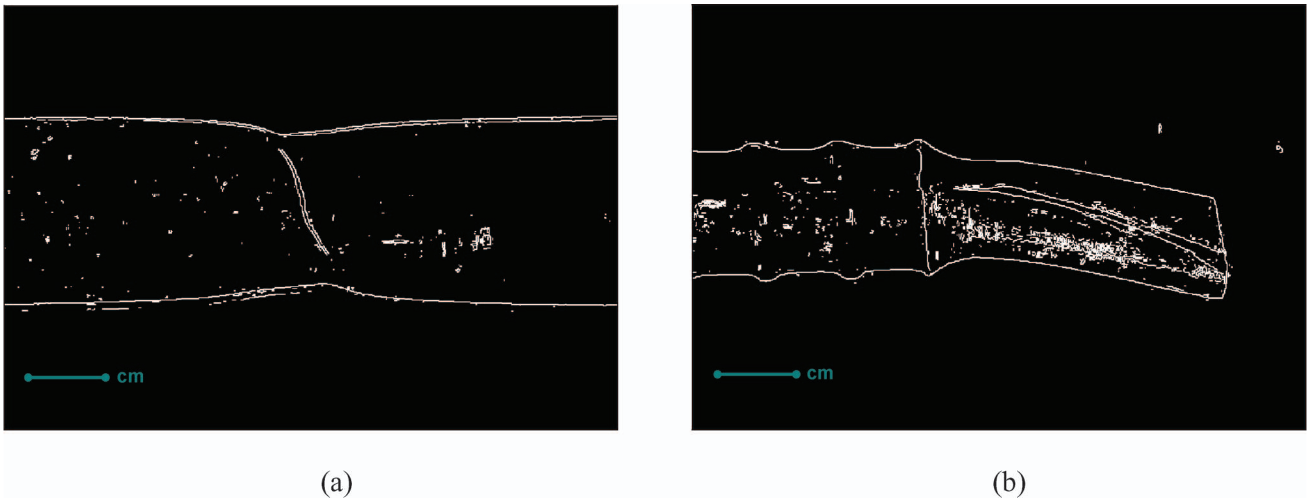


Figure 28. (a) Extracted edges using the Sobel edge detection technique for the steel strip in Figure 16(a); (b) extracted edges using the Sobel edge technique for the steel rebar in Figure 16(b).

projected in four directions (0, 45, 90 and 135°) as features for crack detection in underground pipeline systems.

7. Corrosion detection

There are few papers published on corrosion detection based on image processing techniques alone; however, the capability of image-understanding algorithms is of great interest since they are contactless, nondestructive methods. Because the corrosion process can deteriorate the surface of metals, the corroded surface has a different texture to the rest of the image. Texture can be regarded as the measurement of smoothness, coarseness and regularity (Gonzalez and Woods 1992). Most texture segmentation techniques are based on the pattern recognition concepts described in §4. Pal and Pal (1993) and Reed and Dubuf (1993) provide a comprehensive review of different segmentation techniques. Although the computational cost is high for large windows (Pratt 2001), discrete wavelet transform coefficients are powerful tools to characterise the appropriate features for texture classification as they localise the spatial and frequency characteristics very well (Gunatilake *et al.* 1997).

For subsurface corrosion, a defective area can be recognised based on changes in its surface shape rather than its texture. Stereo cameras are appropriate tools to detect the changes in surface shape, which is useful for detecting the subsurface corrosion. Hanji *et al.* (2003) used this approach to measure the 3D shape of the corroded surface of steel plates. They used a stereoadapter on a regular camera to have the stereovision of the corroded surface. The corresponding regions in the stereo images are detected as high correlative areas in both images. The 3D model of the surface is then reconstructed for measurement purposes. The results of the technique are in good agreement with the measurements of the laser displacement meter (Hanji *et al.* 2003).

Colour is another important attribute of digital image-based corrosion detection. Colour image segmentation surveys are provided by Skarbek and Koschen (1994) and Cheng *et al.* (2001).

Gunatilake *et al.* (1997) used Daubechies (1992) wavelets of order six to detect corrosion areas on aircraft skins. The outcome of their algorithm is a binary image indicating the corroded and non-corroded regions. A three-level wavelet filter bank is used to decompose the image, as described in §5. The low-pass and high-pass filters for this wavelet are shown in Equation (41):

$$h_\varphi = \begin{bmatrix} 0.33267 \\ 0.80689 \\ 0.45988 \\ -0.13501 \\ -0.08544 \\ 0.03523 \end{bmatrix} \quad \text{and} \quad h_\psi = \begin{bmatrix} 0.03523 \\ 0.08544 \\ -0.13501 \\ -0.45988 \\ 0.80689 \\ -0.33267 \end{bmatrix}. \qquad (41)$$

The image is divided into 8×8 pixel blocks. Block-based feature elements lead to a high signal-to-noise ratio and decrease the false detection of corrosion on a surface. Finally, 10 features are assigned to any of these non-overlapping blocks. Each feature is the energy of the block computed from the wavelet coefficients from one of the 10 decomposed frames. The energy of each decomposed frame is defined as the sum of the square of all pixel values belonging to that frame divided by the sum of the square of all pixel values belonging to all decomposed frames of the block. For a better understanding of this definition, a schematic three-level decomposition of an image is shown in Figure 29, where 'L' and 'H' stand for low-pass and high-pass filters respectively.

Each feature can be written mathematically as:

$$f_j(i) = \frac{\sum_{(m,n)\in B(i)} \left(W_j(m,n)\right)^2}{\sum_{j=1}^{10} \sum_{(m,n)\in B(i)} \left(W_j(m,n)\right)^2}, \tag{42}$$

where $f_j\,(i)$ is the j^{th} feature of the i^{th} block, $W_j\,(m,\,n)$ is the wavelet decomposition coefficients of the j^{th} sub-band (as in Figure 29) at $(m,\,n)$ and $B(i)$ is the i^{th} block.

Gunatilake *et al*. (1997) used a nearest neighbour classifier (as described in §4.3.1) to classify corroded regions from corrosion-free areas. The algorithm has a 95% accuracy in detecting corroded regions, as reported by the authors. The authors of the current study have used the same features to automatically segment different textures of an image with an unsupervised classifier. The results are acceptable; however, colour, a very important attribute, is not included in the feature vector described above.

Since corrosion does not always exhibit a repeated texture, and since the lighting in which an image is captured is inconstant, the classical form of texture segmentation described above is incapable of performing reliable defect detection. This problem, along with the aim of integrating several images captured from a single scene to create a corrosion map of a surface, lead to the development of a more sophisticated algorithm that also contains the colour characteristics of an image (Siegel and Gunatilake 1998).

In their new algorithm, Siegel and Gunatilake (1998) converted the RGB image into a more uncorrelated colour space of YIQ, where Y represents luminance, and I and Q represent chrominance information. Equation (43) shows how RGB and YIQ colour components are related (Buchsbaum 1968, Pratt 2001):

$$\begin{pmatrix} Y \\ I \\ Q \end{pmatrix} = \begin{pmatrix} 0.29890 & 0.58660 & 0.11448 \\ 0.59598 & -0.27418 & -0.32180 \\ 0.21147 & -0.52260 & 0.31110 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}. \tag{43}$$

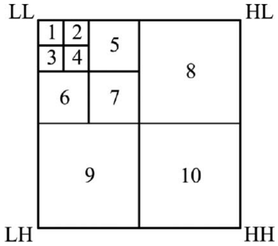


Figure 29. Three-level wavelet decomposition notation used by Gunatilake *et al*. (1997).

The Battle–Lemaire (BL) (Strang and Nguyen 1996) wavelet transform filter is used to obtain a three-level decomposition of the Y component without down-sampling at each stage. Consequently, 10 equal-sized images are obtained as shown in Figure 30. The two-level decomposition of the I and Q components are computed using the normal wavelet transform with down-sampling at each stage. The Y image is divided into 32 × 32 non-overlapping pixel blocks; for each of these blocks, 10 features from the Y decomposition image and four features from the I and Q decomposition images are extracted. Each of the first nine features derived from the Y image is calculated as:

$$f^{SB_k}(x,y) = \frac{\sum_{i=-\frac{b}{2}}^{\frac{b}{2}} \sum_{j=-\frac{b}{2}}^{\frac{b}{2}} wSB_k^2}{\sum_{i=-\frac{b}{2}}^{\frac{b}{2}} \sum_{j=-\frac{b}{2}}^{\frac{b}{2}} w[LH_k^2 + HH_k^2 + HL_k^2]}, \tag{44}$$

where

$$w = w(x+i, y+j), \tag{45}$$

$$SB_k = SB_k(x+i, y+j), \tag{46}$$

$$LH_k = LH_k(x+i, y+j), \tag{47}$$

$$HH_k = HH_k(x+i, y+j), \tag{48}$$

and

$$HL_k = HL_k(x+i, y+j). \tag{49}$$

Equation (44) is the ratio of a detailed image, resulting from decomposition and the total energy of all details in that level of decomposition. In this equation, *HH, HL* and *LH* are the decomposed images, as shown in Figure 30. The coordinate $(x,\,y)$ is the centre of the blocks in the Y decomposed images, b is the size of the block, SB_k is either *HH, HL*, or *LH* at level k of the decomposition, and w is a Gaussian weighted mask as described in §2. In fact, the computation of these nine features is similar to the traditional texture segmentation techniques using wavelet transform where a Gaussian mask is used to define the weighting factor to calculate the energy of each block.

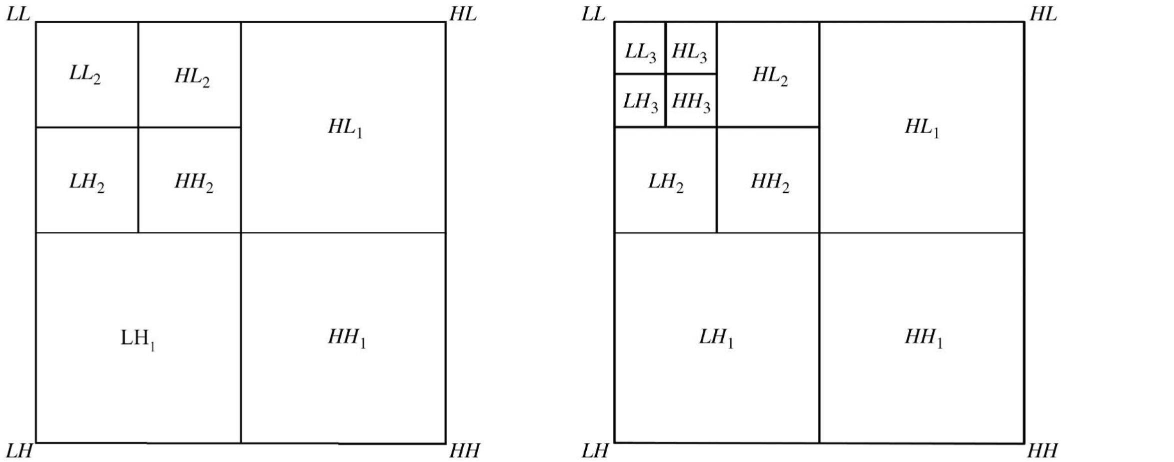


Figure 30. (a) Two-level wavelet decomposition of an image, and (b) three-level wavelet decomposition of an image (notation used by Siegel and Gunatilake (1998)).

The 10th feature that is extracted from image Y is:

$$f^{LL_3}(x,y) = \frac{\sum_{k=1}^{3}\sum_{i=-\frac{b}{2}}^{\frac{b}{2}}\sum_{j=-\frac{b}{2}}^{\frac{b}{2}} w[LH_k^2 + HH_k^2 + HL_k^2]}{\sum_{i=-\frac{b}{2}}^{\frac{b}{2}}\sum_{j=-\frac{b}{2}}^{\frac{b}{2}} wLL_3^2}, \tag{50}$$

where b is the size of the block, SB_k is either HH, HL or LH at level k of the decomposition and w is a Gaussian weighted mask. The parameters w, SB_k, LH_k, HH_k, and HL_k are defined by Equations (45), (46), (47), (48) and (49), respectively. Equation (50) represents the ratio of the whole energy of all the details and the approximation energy of the Y image after a three-level wavelet decomposition.

Four more features are extracted from the two-level wavelet decompositions of the I and Q image components, as shown in Equations (51) and (52):

$$f_k^I(x_k', y_k') = \frac{\sum_{i=-\frac{b}{2}}^{\frac{b}{2}}\sum_{j=-\frac{b}{2}}^{\frac{b}{2}} w[LH_k^2 + H_k^2 + HL_k^2]}{\sum_{i=-\frac{b}{2}}^{\frac{b}{2}}\sum_{j=-\frac{b}{2}}^{\frac{b}{2}} w\mathrm{LL}_2^2}, \tag{51}$$

and

$$f_k^Q(x_k', y_k') = \frac{\sum_{i=-\frac{b}{2}}^{\frac{b}{2}}\sum_{j=-\frac{b}{2}}^{\frac{b}{2}} w[LH_k^2 + H_k^2 + HL_k^2]}{\sum_{i=-\frac{b}{2}}^{\frac{b}{2}}\sum_{j=-\frac{b}{2}}^{\frac{b}{2}} w\mathrm{LL}_2^2}, \tag{52}$$

where

$$w = w(x_k' + i, y_k' + j), \tag{53}$$

$$SB_k = SB_k(x_k' + i, y_k' + j), \tag{54}$$

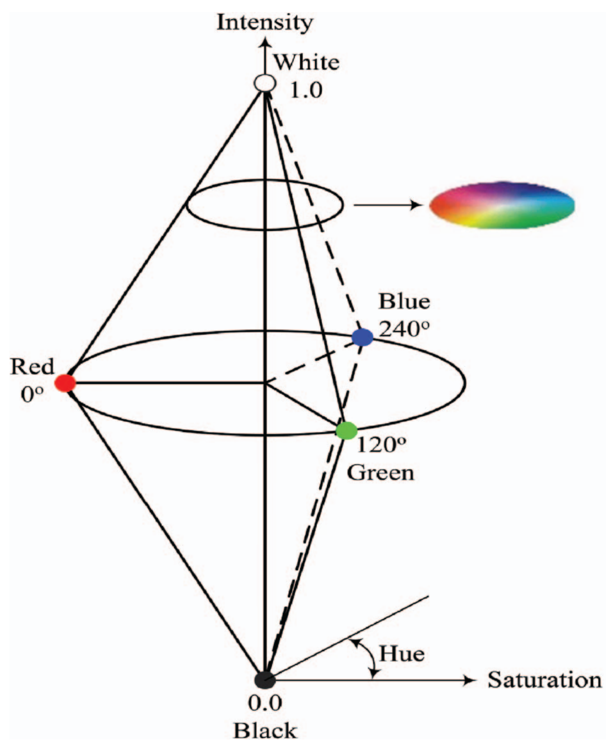


Figure 31. HSI colour space.

$$LH_k = LH_k(x_k' + i, y_k' + j), \tag{55}$$

$$HH_k = HH_k(x_k' + i, y_k' + j), \tag{56}$$

and

$$HL_k = HL_k(x_k' + i, y_k' + j), \tag{57}$$

where b is the size of the block, SB_k is either HH, HL, or LH at level k of the decomposition, w is a Gaussian

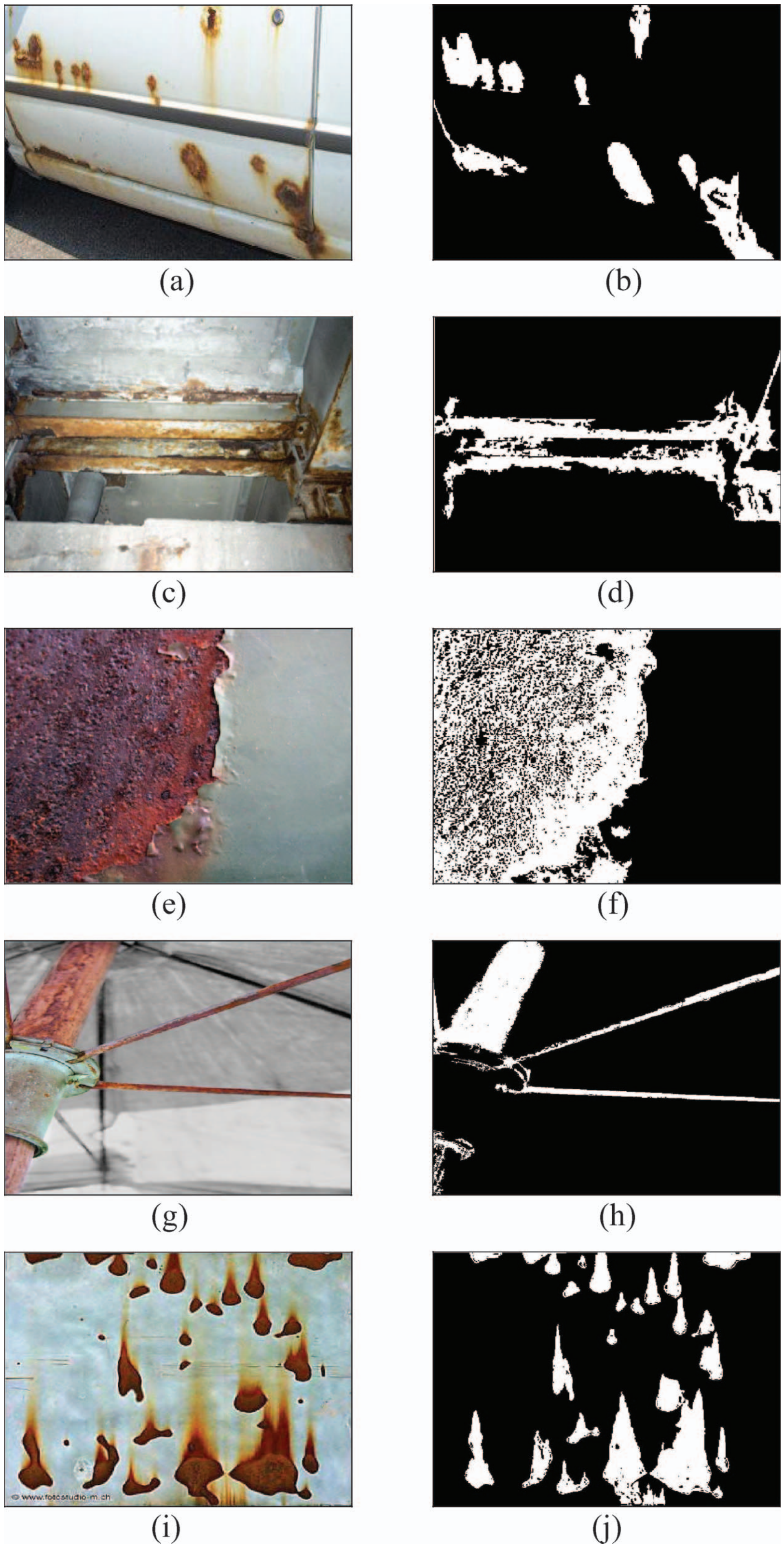


Figure 32. The result of binarising the saturation component of original images (a), (c), (e), (g) and (i), in order to extract the corroded area of each image, is presented in images (b), (d), (f), (h) and (j), respectively.

weighted mask, k is the level of decomposition (here $k = 1, 2$) and (x'_k, y'_k) in I and Q is the corresponding coordinate of the block centre (x, y) in Y at the k^{th} level of decomposition. This can be computed as:

$$(x'_k, y'_k) = \left(\frac{x}{2^k}, \frac{y}{2^k}\right). \tag{58}$$

After computing (x'_k, y'_k), a 32×32 window centred at (x'_k, y'_k) is selected to calculate the feature values, as shown in Equations (51) and (52).

The 32×32 windows in I and in Q will overlap as the level of decomposition increases. For the higher levels of decomposition, this process will result in a better estimation of the low-frequency signals in the chrominance characteristics of the image. Equations (51) and (52) are ratios of the total energy of the details in the k^{th} level of decomposition and the energy of the second level approximation for the I and Q components. After extracting the described features, Siegel and Gunatilake (1998) used a feed-forward neural network consisting of 14 input, 40 hidden and 2 output neurons. The possible outputs of the algorithm are: corrosion with high confidence, or a corrosion with low confidence or a corrosion-free region. The decision making function is based on the two outputs of the neural network and a threshold T that is experimentally selected as 0.65. The confidence is defined as the absolute value of the difference between the two outputs:

$$\begin{cases} \left\{\begin{array}{c} \text{output}(1) > \text{output}(2) \\ \& \\ \text{confidence} \geq T \end{array}\right\} \Rightarrow \begin{array}{l} \text{corrosion} \\ \text{(high confidence)} \end{array} \\ \left\{\begin{array}{c} \text{output}(1) > \text{output}(2) \\ \& \\ \text{confidence} < T \end{array}\right\} \Rightarrow \begin{array}{l} \text{corrosion} \\ \text{(low confidence)} \end{array} \\ \text{output}(1) < \text{output}(2) \Rightarrow \text{corrosion}-\text{free} \end{cases} \tag{59}$$

Because of the above decision making procedure, it is possible to process multiple images and perform information fusion over captured data. In addition, a single corrosion map can be generated where each region has the largest confidence value extracted from different images. The probability of correct detection for this algorithm is 94% (Siegel and Gunatilake 1998).

Another way to evaluate the colour characteristics of a corroded region is to convert the RGB colour image into the HSI colour space. It is possible to express the colour characteristics independent of brightness in the HSI colour space. For this reason, the HSI colour space is a suitable choice for identifying corroded areas quantitatively (Choi and Kim 2005). The following equations show how the RGB components of an image can be converted to the HSI components:

$$I = \frac{1}{3}(\mathrm{R} + \mathrm{G} + \mathrm{B}), \tag{60}$$

$$S = 1 - \left(\frac{3}{R + G + B}\right)(\min(\mathrm{R}, \mathrm{G}, \mathrm{B})), \tag{61}$$

and

$$H = \cos^{-1} \frac{\frac{1}{2}[(\mathrm{R} - \mathrm{G}) + (\mathrm{R} - \mathrm{B})]}{\sqrt{(\mathrm{R} - \mathrm{G})^2 + (\mathrm{R} - \mathrm{B})(\mathrm{G} - \mathrm{B})}}. \tag{62}$$

Figure 31 shows the HSI colour space in which the saturation varies from 0 to 1, the hue varies from 0 to 360°, and the intensity varies from 0 (black) to 1 (white).

Binarising the saturation component of an image will segment the pixels that have higher saturation values than the rest of the image, which, in many cases, can result in segmenting the corroded area. This indicates the significance of the saturation component. Figure 32 shows the result of binarising the saturation component in actual corrosion images using the threshold technique described in §6.1, where the white areas in Figures 32b, d, f, h and j are the potential corroded areas.

Choi and Kim (2005) classified several types of corrosion defects. They proposed to divide each H and S component into 10×10 pixel blocks, and then treat the histogram of each block like a distribution of random variables. After applying the PCA and varimax approach, it was concluded that the mean H value, the mean S value, the median S value, the skews of the S distribution and the skews of the I distribution are appropriate features to be assigned to each block for classification (Choi and Kim 2005). The co-occurrence matrix is used for texture feature extraction based on the azimuth difference of points on a surface. This approach may not be useful for the problem in question since it requires microscopes to capture the images; the magnification factor of the tested images by Choi and Kim (2005) is between 50 and 500, which is far beyond the magnification factor of regular digital cameras.

8. Summary and conclusions

Among the possible techniques for inspecting civil infrastructure, the use of optical instrumentation that relies on image processing is a less time consuming and inexpensive alternative to current monitoring methods. This paper provides a survey and an evaluation of some of the promising vision-based approaches for automatic detection of missing (deformed) members, cracks and corrosion in civil infrastructure systems. Several examples that are based on laboratory studies are presented in the paper to illustrate the utility, as well as the limitations, of the leading approaches.

Image registration can be the key to resolving issues such as visual detection of missing or deformed structural members and 3D reconstruction of truss systems, which can lead to localising the detected changes. Different feature detection and feature matching techniques for image registration are presented in this paper, including the Moravec operator, the Harris (Plessey) operator, the SIFT detector, and the RANSAC algorithm; however, current registration techniques need to be improved in order to reconstruct a 3D model of truss-like structures for localisation purposes.

Morphological image processing techniques are promising approaches to segment the probable crack-like patterns from the rest of the image. Two morphological algorithms based on modifying the top-hat and bottom-hat transformations are the best algorithms for this purpose, as verified by the authors. Several edge-based techniques are discussed, including Sobel, Roberts, Prewitt, Canny, Laplacian, Laplacian of Gaussian, FFT and wavelet transform approaches. Crack detection techniques discussed in this paper are highly promising, but the problem of irregular backgrounds requires further study.

Discrete wavelet filter banks are introduced and several applications of wavelet transform coefficients in edge detection, crack detection, image enhancement and compression, as well as corrosion detection, are described.

In order to segment corrosion-like regions from the rest of the image, both texture and colour analysis should be applied. Multi-resolution wavelet analysis is a powerful tool to characterise the appropriate features for texture classification. Decision making is a vital step in detecting cracks and corrosion regions, as well as the texture classification. The concept of pattern recognition, including supervised and unsupervised classification, and the performances of some of these classification techniques such as k-nearest neighbour, k-means and neural network are discussed. Neural network classification has an acceptable performance in most cases. To perform colour analysis, YIQ and HSI colour spaces appear to provide suitable features for the classification process. Corrosion detection requires more research to correctly segment and classify the defected regions.

Neuro-fuzzy systems simultaneously benefit from the data imprecision tolerance (vague definitions) of fuzzy logic systems and the tolerance of the neural networks for noisy data. The easily comprehendible linguistic terms and if-then rules of the fuzzy systems and the learning capabilities of the neural networks are fused into a neuro-fuzzy system. These capabilities make the neuro-fuzzy expert systems appropriate for pattern recognition and defect classification.

Data fusion and image acquisition of a scene using several cameras or different lighting conditions are very interesting aspects of the class of problems under discussion, and which require more research. The results discussed within this paper show that using image processing and pattern recognition techniques are promising approaches for contactless nondestructive health monitoring of many civil infrastructures.

Acknowledgements

This study was supported in part by grants from the National Science Foundation.

References

Abdel-Qader, I., Abudayyeh, O., and Kelly, M.E., 2003. Analysis of edge-detection techniques for crack identification in bridges. *Journal of Computing in Civil Engineering*, 17 (4), 255–263.

Abdou, I., 1973. *Quantitative methods of edge detection*. Los Angeles, CA: Image Processing Institute, University of Southern California. Technical Report USCIPI Report 830.

Abdou, I.E. and Pratt, W.K., 1979. Quantitative design and evaluation of enhancement/thresholding edge detectors. *Proceedings of IEEE*, 67 (5), 753–763.

Abtoine, J.-P., Murenzi, R., Vandergheynst, P., and Ali, S.T., 2004. *Two-dimensional wavelets and their relatives*. Cambridge, UK: Cambridge University Press.

Achler, O. and Trivedi, M.M., 2004. Camera based vehicle detection, tracking, and wheel baseline estimation approach. *In*: *Proceedings of IEEE conference on intelligent transportation systems*. ITSC, 743–748.

Al-Otum, H.M., 2003. Morphological operators for color image processing based on mahalanobis distance measure. *Optical Engineering*, 42 (9), 2595–2606.

Alageel, K. and Abdel-Qader, I., 2002. Harr transform use in image processing. Kalamazoo, MI: Department of Electrical and Computer Engineering, Western Michigan University, Technical Report.

Bachmann, G., Narici, L., and Beckenstein, E., 2000. *Fourier and wavelet analysis*. New York, NY: Springer.

Benning, W., Görtz, S., Lange, J., Schwermann, R., and Chudoba, R., 2003. Development of an algorithm for automatic analysis of deformation of reinforced concrete structures using photogrammetry. *VDI Berichte*, 1757, 411–418.

Bosc, M., Heitz, F., Armspach, J.-P., Namer, I., Gounot, D., and Rumbachc, L., 2003. Automatic change detection in multimodal serial MRI: application to multiple sclerosis lesion evolution. *Neuroimage*, 20, 643–656.

Brown, L.G., 1992. A survey of image registration techniques. *ACM Computing Surveys*, 24 (4), 325–376.

Bruzzone, L. and Serpico, S.B., 1997. An iterative technique for the detection of land-cover transitions in multitemporal remote-sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 35 (4), 858–867.

Buchsbaum, W.H., 1968. *Color TV servicing*. 2nd ed. Englewood, Cliffs: Prentice Hall Press.

Canny, J., 1986. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8 (6), 679–698.

Chae, M.J., 2001. *Automated interpretation and assessment of sewer pipeline*. Thesis (PhD). Purdue University.

Chang, S.G., Yu, B., and Vetterli, M., 2000. Spatially adaptive wavelet thresholding with context modeling for image denoising. *IEEE Transactions on Image Processing*, 9 (9), 1522–1531.

Cheng, H.D., Jiang, X.H., Sun, Y., and Wang, J., 2001. Color image segmentation: advances and prospects. *Pattern Recognition*, 34 (12), 2259–2281.

Choi, K. and Kim, S., 2005. Morphological analysis and classification of types of surface corrosion damage by digital image processing. *Corrosion Science*, 47 (1), 1–15.

Chung, H.-C., Liangand, J., Kushiyama, S., and Shinozuka, M., 2004. Digital image processing for non-linear system identification. *International Journal of Non-linear mechanics*, 39, 691–707.

Coffey, J.M., 1988. Non-destructive testing – the technology of measuring defects. *CEGB Research*, 21, 36–47.

Collins, R.T., Lipton, A.J., and Kanade, T., 2000. Introduction to the special section on video surveillance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22 (8), 745–746.

Comer, M.L. and Delp, E.J., 1999. Morphological operations for color image processing. *Journal of Electronic Imaging*, 8 (3), 279–289.

Coppin, P.R. and Bauer, M.E., 1996. Digital change detection in forest ecosystems with remote sensing imagery. *Remote Sensing Reviews*, 13 (3–4), 207–234.

Daubechies, I., 1992. *Ten lectures on wavelets*. Philadelphia, USA: SIAM.

Davis, L.S., 1975. A survey of edge detection techniques. *Computer Graphics and Image Processing*, 4, 248–270.

Deer, P. and Eklund, P., 2002. Values for the fuzzy c-means classifier in change detection for remote sensing. *In*: *Proceedings of 9th international conference on information processing and management of uncertainty*, 187–194.

DeVore, R.A., Jawerth, B., and Lucier, B.J., 1992. Image compression through wavelet transform coding. *IEEE Transactions on Information Theory*, 38 (2), 719–746.

Duda, R.O. and Hart, P.E., 1973. *Pattern classification and scene analysis*. New York, NY: John Wiley & Sons.

Duda, R.O., Hart, P.E., and Stork, D.G., 2001. *Pattern classification*. 2nd ed. New York, NY: Wiley.

Dudziak, M.J., Chervonenkis, A.Y., and Chinarov, V., 1999. Nondestructive evaluation for crack, corrosion, and stress detection for metal assemblies and structures. *In*: *Proceedings of SPIE – nondestructive evaluation of aging aircraft, airports and aerospace hardware III*, 3586, 20–31.

Dumskyj, M.J., Aldington, S.J., Dore, C.J., and Kohner, E.M., 1996. The accurate assessment of changes in retinal vessel diameter using multiple frame electrocardiograph synchronised fundus photography. *Current Eye Research*, 15 (6), 625–632.

Edgington, D.R., Salamy, K.A., Risi, M., Sherlock, R.E., Walther, D., and Koch, C., 2003. Automated event detection in underwater video. *Oceans Conference Record (IEEE)*, 5, 2749–2753.

Faugeras, O., 1993. *Three-dimensional computer vision: a geometric view point*. Cambridge, MA: MIT Press.

Fischler, M.A. and Bolles, R.C., 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24 (6), 381–395.

Frei, W. and Chen, C., 1977. Fast boundary detection: a generalization and a new algorithm. *IEEE Transactions on Computers*, C-26 (10), 988–998.

Garcia-Alegre, M.C., Ribeiro, A., Guinea, D., and Cristobal, G., 2000. Eggshell defects detection based on color processing. *Machine Vision Applications in Industrial Inspection VIII*, 3966, 280–287.

Giakoumis, I., Nikolaidis, N., and Pitas, I., 2006. Digital image processing techniques for the detection and removal of cracks in digitized paintings. *IEEE Transactions on Image Processing*, 15 (1), 178–188.

Goldin, S.E. and Rudahl, K.T., 1986. Tutorial image processing system: a tool for remote sensing training. *International Journal of Remote Sensing*, 7 (10), 1341–1348.

Gonzalez, R.C. and Wintz, P., 1987. *Digital image processing*. 2nd ed. Boston, MA: Addison-Wesley.

Gonzalez, R.C. and Woods, R.E., 1992. *Digital image processing*. Boston, MA: Addison-Wesley.

Gonzalez, R.C., Woods, R.E., and Eddins, S.L., 2004. *Digital image processing using MATLAB*. Upper Saddle River, NJ: Prentice Hall.

Gunatilake, P., Siegel, M.W., Jordan, A.G., and Podnar, G.W., 1997. Image understanding algorithms for remote visual inspection of aircraft surfaces. *In*: *Proceedings of SPIE – international society for optical engineering*, 3027, 2–13.

Hanji, T., Tateishi, K., and Kitagawa, K., 2003. 3-D shape measurement of corroded surface by using digital stereography. *In*: *Proceedings of 1st international conference on structural health monitoring and intelligent infrastructure*, 1, 699–704.

Harris, C. and Stephens, M., 1988. A combined corner and edge detector. *In: Alvey vision conference*, Plessey Research Roke Manor, UK, 147-152.

Hartley, R. and Zisserman, A., 2000. *Multiple view geometry in computer vision*. Cambridge, UK: Cambridge University Press.

Henstock, P.V. and Chelberg, D.M., 1996. Automatic gradient threshold determination for edge detection. *IEEE Transactions on Image Processing*, 5 (5), 784–787.

Ho, S.K., White, R.M., and Lucas, J., 1990. Vision system for automated crack detection in welds. *Measurement Science and Technology*, 1 (3), 287–294.

Hogg, D.C., 1993. Shape in machine vision. *Image and Vision Computing*, 11 (6), 309–316.

Horn, B.K.P., 1986. *Robot vision*. Cambridge, MA: MIT Press, McGraw-Hill Book Co.

Huertas, A. and Nevatia, R., 2000. Detecting changes in aerial views of man-made structures. *Image and Vision Computing*, 18, 583–596.

Jolliffe, I.T., 2002. *Principal component analysis*. 2nd ed. New York, NY: Springer.

Kaseko, M.S., Lo, Z.-P., and Ritchie, S.G., 1994. Comparison of traditional and neural classifiers for pavement-crack detection. *Journal of Transportation Engineering*, 120 (4), 552–569.

Kumar, S. and Taheri, F., 2007. Neuro-fuzzy approaches for pipeline condition assessment. *Nondestructive Testing and Evaluation*, 22 (1), 35–60.

Lebart, K., Trucco, E., and Lane, D., 2000. Real-time automatic sea-floor change detection from video. *Oceans Conference Record (IEEE)*, 2, 1337–1343.

Lee, K.H., 2005. *First course on fuzzy theory and applications*. Berlin: Springer-Verlag.

Lemieux, L., Wieshmann, U.C., Moran, N.F., Fish, D.R., and Shorvon, S.D., 1998. The detection and significance of subtle changes in mixed-signal brain lesions by serial MRI scan matching and spatial normalization. *Medical Image Analysis*, 2 (3), 227–242.

Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60 (2), 91–110.

Lu, D., Mausel, P., Brondizio, E., and Moran, E., 2004. Change detection techniques. *International Journal of Remote Sensing*, 25 (12), 2365–2407.

Mallat, S.G., 1989. Multifrequency channel decompositions of images and wavelet models. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 37 (12), 2091–2110.

Mallat, S. and Zhong, S., 1992. Characterization of signals from multiscale edges. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14 (7), 710–732.

Matheron, G., 1975. *Random sets and integral geometry*. New York, NY: Wiley.

Meyer, F., 1986. Automatic screening of cytological specimens. *Computer Vision, Graphics and Image Processing*, 35 (3), 356–369.

Minkowski, H., 1903. Volumen und oberfläche. *Mathematische Annalen*, 57 (4), 447–495.

Misiti, M., Misiti, Y., Oppenheim, G., and Poggi, J.-M., 2006. *Wavelet toolbox users guide,* version 3 [online]. Natick, MA: MathWorks, Inc. Available from: http://www.mathworks.com/access/helpdesk/help/pdf_doc/wavelet/wavelet_ug.pdf

Moravec, H.P., 1977. Towards automatic visual obstacle avoidance. *In*: *Proceedings of 5th international joint conference on artificial intelligence*, 584.

Moravec, H.P., 1979. Visual mapping by a robot rover. *In*: *Proceedings of international joint conference on artificial intelligence*, 598–600.

Moselhi, O. and Shehab-Eldeen, T., 2000. Classification of defects in sewer pipes using neural networks. *Journal of Infrastructure Systems*, 6 (3), 97–104.

Nieniewski, M., Chmielewski, L., Jozwik, A., and Sklodowski, M., 1999. Morphological detection and feature-based classification of cracked regions in ferrites. *Machine Graphics and Vision*, 8 (4), 699–712.

Pal, N.R. and Pal, S.K., 1993. A review of image segmentation techniques. *Pattern Recognition*, 26 (9), 1277–1294.

Parks, D. and Gravel, J.-P., 2005. *Corner detectors* [online]. Center for Intelligent Machines, McGill University. Available from: http://www.cim.mcgill.ca/~dparks/CornerDetector/index.htm [Accessed 1 December 2005].

Peng, L., Zhao, Z., Cui, L., and Wang, L., 2004. Remote sensing study based on irsa remote sensing image processing system. *In*: *Proceedings of IEEE international geoscience and remote sensing symposium: science for society: exploring and managing a changing planet (IGARSS)*, 7, 4829–4832.

Pines, D. and Aktan, A.E., 2002. Status of structural health monitoring of long-span bridges in the united states. *Progress in Structural Engineering and Materials*, 4, 372–380.

Porter, R. and Canagarajah, N., 1996. A robust automatic clustering scheme for image segmentation with wavelets. *IEEE Transactions on Image Processing*, 5 (4), 662–665.

Poudel, U.P., Fu, G., and Ye, J., 2005. Structural damage detection using digital video imaging technique and wavelet transformation. *Journal of Sound and Vibration*, 286 (4–5), 869–895.

Prasad, L. and Iyenger, S., 1997. *Wavelet analysis with applications to image processing*. Boca Raton, FL: CRC Press.

Pratt, W.K., 2001. *Digital image processing*. 3rd ed. New York, NY: Wiley.

Prewitt, J.M.S., 1970. *Object enhancement and extraction*. New York, NY: Academic Press.

Radke, R.J., Andra, S., Al-Kofahi, O., and Roysam, B., 2005. Image change detection algorithms: a systematic survey. *IEEE Transactions on Image Processing*, 14 (3), 294–307.

Reed, T.R. and Dubuf, J.M.H., 1993. A review of recent texture segmentation and feature extraction techniques. *CVGIP: Image Understanding*, 57 (3), 359–372.

Rey, D., Subsol, G., Delingette, H., and Ayache, N., 2002. Automatic detection and segmentation of evolving processes in 3D medical images: application to multiple sclerosis. *Medical Image Analysis*, 6 (2), 163–179.

Ringer, M. and Morris, R.D., 2001. Robust automatic feature detection and matching between multiple images. Research Institute for Advanced Computer Science, RIACS Technical Report 01.27.

Roberts, L.G., 1965. *Machine perception of three-dimensional solids*. Cambridge, MA: MIT Press, 159–197.

Rousseeuw, P.J., 1987. *Robust registration and outlier detection*. New York, NY: John Wiley & Sons.

Salembier, P., 1990. Comparison of some morphological segmentation algorithms based on contrast enhancement. application to automatic defect detection. *In*: *Proceedings of 5th European signal processing conference*, 833–836.

Schlag, J.F., Sanderson, C., Neuman, C.P. and Wimberly, F.C., 1983. Implementation of automatic focusing algorithms for a computer vision system with camera. Carnegie-Mellon University, Pittsburgh, PA 15213, Technical Report.

Scott, G. and Longuet–Higgis, H., 1991. An algorithm for associating the features of two patterns. *In*: *Proceedings Royal Society London B*, 244, 21–26.

Serra, J., 1982. *Image analysis and mathematical morphology*. London, UK: Academic Press.

Shinozuka, M., 2003. Homeland security and safety. *In*: *Proceedings of structural health monitoring and intellegent infrastructure*, 2, 1139–1145.

Shinozuka, M., Chung, H.-C., Ichitsubo, M., and Liang, J., 2001. System identification by video image processing. *In*: *Proceedings of SPIE – international society for optical engineering*, 4330, 97–107.

Shubinsky, G., 1994. *Visual and infrared imaging for bridge inspection*. Northwestern University Basic Industrial Research Laboratory.

Siegel, M. and Gunatilake, P., 1998. Remote enhanced visual inspection of aircraft by a mobile robot. *IEEE workshop on emerging technologies, intelligent measurement and virtual systems for instrumentation and measurement*. St. Paul, MN, USA.

Singh, A., 1989. Digital change detection techniques using remotely-sensed data. *International Journal of Remote Sensing*, 10 (6), 989–1003.

Sinha, S.K., 2000. Automated underground pipe inspection using a unified image processing and artificial intelligence methodology. Thesis (PhD). University of Waterloo, Waterloo, Ontario, Canada.

Sinha, S.K. and Fieguth, P.W., 2006a. Classification of underground pipe scanned images using feature extraction and neuro-fuzzy algorithm. *Automation in Construction*, 15 (1), 58–72.

Sinha, S.K. and Fieguth, P.W., 2006b. Morphological segmentation and classification of underground pipe images. *Machine Vision and Applications*, 17 (1), 21–31.

Sinha, S.K., Fieguth, P.W., and Polak, M.A., 2003. Computer vision techniques for automatic structural assessment of underground pipes. *Computer Aided Civil and Infrastructure Engineering*, 18 (2), 95–112.

Skarbek, W. and Koschen, A., 1994. Colour image segmentation: a survey. Institute for Technical Informatics, Technical University of Berlin, Technical Report.

Strang, G. and Nguyen, T., 1996. *Wavelets and filter banks*. Wellesley, MA: Wellesley–Cambridge Press.

Thirion, J.-P. and Calmon, G., 1999. Deformation analysis to detect and quantify active lesions inthree-dimensional medical image sequences. *IEEE Transactions on Medical Imaging*, 18 (5), 429–441.

Torr, P.H.S. and Murray, D.W., 1997. The development and comparison of robust methods for estimating the fundamental matrix. *International Journal of Computer Vision*, 24 (3), 271–300.

Tsao, S., Kehtarnavaz, N., Chan, P., and Lytton, R., 1994. Image-based expert-system approach to distress detection on CRC pavement. *Journal of Transportation Engineering*, 120 (1), 62–64.

Villasenor, J.D., Belzer, B., and Liao, J., 1995. Wavelet filter evaluation for image compression. *IEEE Transactions on Image Processing*, 4 (8), 1053–1060.

Wang, K.C., Nallamothu, S., and Elliott, R.P., 1998. Classification of pavement surface distress with an embedded neural net chip. *In*: *Artificial neural networks for civil engineers: advanced features and applications,* 131–161.

Watanabe, S., Miyajima, K., and Mukawa, N., 1998. Detecting changes of buildings from aerial images using shadow and shading model. *In*: *Proceedings of 14th international conference on pattern recognition*, 2 (2), 1408–1412.

Yu, M., Wang, R., Jiang, G., Liu, X., and Cho, T.-Y., 2004. New morphological operators for color image processing. *In*: *Proceedings of IEEE region 10 annual international conference*, A, 443–446.

Zhang, Z., 1998. Determining the epipolar geometry and its uncertainty: a review. *International Journal of Computer Vision*, 27 (2), 161–195.

Ziou, D. and Tabbone, S., 1998. Edge detection techniques – an overview. *Pattern Recognition and Image Analysis*, 8 (4), 537–559.

Zitová, B. and Flusser, J., 2003. Image registration methods: a survey. *Image and Vision Computing*, 21 (11), 977–1000.