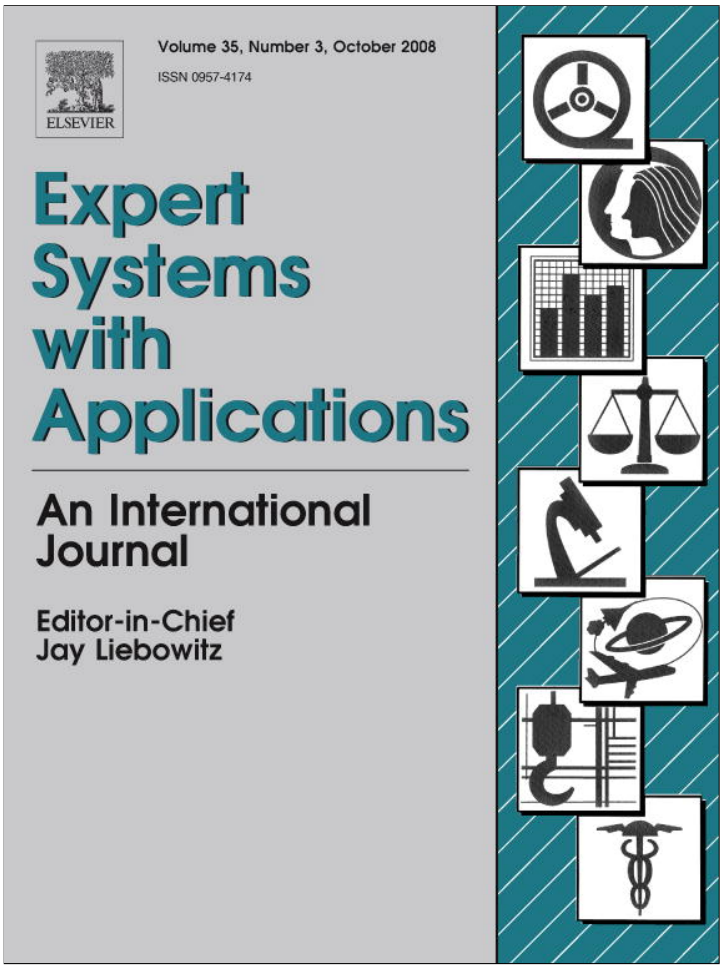


Available online at www.sciencedirect.com
ScienceDirect
Expert Systems with Applications 35 (2008) 1327–1337
Expert Systems with Applications
www.elsevier.com/locate/eswa

Automated diagnosis of sewer pipe defects based on machine learning approaches

Ming-Der Yang *, Tung-Ching Su

Department of Civil Engineering, National Chung Hsing University, 250, Kuo Kuang Road., Taichung 402, Taiwan

Abstract

In sewage rehabilitation planning, closed circuit television (CCTV) systems are the widely used inspection tools in assessing sewage structural conditions for non man entry pipes. Currently, the assessment of sewage structural conditions by manually interpretation on CCTV images seems inefficient, especially for several thousands of frames in one inspection plan. Also, the assessment work significantly involves engineers' eye sight and professional experience. With a purpose of assisting general staffs in diagnosing pipe defects on CCTV inspection images, a diagnostic system by applying machine learning approaches is proposed in this paper. This research was first to use image process techniques, including wavelet transform and computation of co-occurrence matrices, for describing the textures of the pipe defects. Then, three neural network approaches, back-propagation neural network (BPN), radial basis network (RBN), and support vector machine (SVM), were adopted to classify pipe defect patterns, and their performances were compared and discussed. The diagnostic system of pipe defects was applied to a sewer system in the 9th district, Taichung City which is the largest city in middle Taiwan. The result shows that the diagnosis accuracy of 60% derived by SVM is the best and also better than the diagnosis accuracy of 57.4% derived by a Bayesian classifier.

Keywords: CCTV images; Sewer pipe defects; Diagnostic system; Textural features

1. Introduction

Sewage rehabilitation plays an important role as sewage construction but is not easily processed due to uncertainty of occurrence of sewer pipe defects. Thus, worldwide engineers pay a greater attention to the proactive and preventive repair strategies than the traditional approach of regular sewer maintenance (Fenner, 2000). Before undertaking sewer rehabilitation, four major series steps of sewer rehabilitation planning including inspection of sewage, assessment of sewage structural conditions, computation of structural condition grades, and determination of rehabilitation methods and substitution materials have to be finished (Yang & Su, 2006, 2007). Various tools or technologies such as closed circuit television (CCTV) cameras mounted on robots, ground piercing radar (GPR), sonar and infrared thermograph, are developed to assist engineers in sewage inspection (Cordes, Berns, Eberl, Ilg, & Suna, 1997; Fenner, 2000; Makar, 1999; Moselhi & Shehab-Eldeen, 1999; Wirahadikusumah, Abraham, Iseley, & Prasanth, 1998). CCTV, one of the most popular inspection tools because of its commercial availability, is usually mounted on robot to be putted inside sewer pipes from a manhole and remotely-controlled outside to acquire images of inner pipe (Makar, 1999). In addition, the advantages of mobile CCTV system include fewer inspectors needed, more safety-ensured to inspectors, and more detailed data of distance and slope possibly recorded (Madryas & Przybyla, 1998).

Traditionally, pipe defects are generally diagnosed by human interpretation on CCTV inspection images. It remains to improve the technology of interpreting an enormous quantity of CCTV inspection images due to human fatigue, expertise-dependence, and inefficiency. To

* Corresponding author. Tel.: +886 4 2284 0440x214; fax: +886 4 2286 2857.
E-mail address: mdyang@dragon.nchu.edu.tw (M.-D. Yang).

0957-4174/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2007.08.013

overcome these limitations, some computer vision systems based on morphologies or geometries of pipe defects have been tried (McKim & Sinha, 1999; Moselhi & Shehab-Eldeen, 1999; Xu, Luxmoore, & Davies, 1998). Sinha and Fieguth (2006) used mathematical morphology methodology for segmentation of pipe defects including crack/hole, joint, collapse pipe, and lateral, but the method was difficult to classify those defect patterns into various classes based on their severity of defects. It was suggested that the segmented objects have to be further processed to be classified according to the severity of defects by using other shape or textural features. This paper attempts to apply machine learning approaches to develop a diagnostic system with better capacity in diagnosing pipe defects based on textural features instead of morphological or geometrical features.

Like a fingerprint databases, the straightforward approach is to establish a complete databank, which stores an enormous amount of CCTV inspection images collected from many inspection cases. Any inspection image can be classified into one category of sewage structural conditions by comparing the defect characteristic with the stored ones. However, it is almost impossible to establish the databank due to the natural shape irregularities of sewer pipe defects and the complex imaging environment. Therefore, this paper attempts to extract specific principal textural features of pipe defects from CCTV inspection images acquired in one inspection mission to train the diagnostic system and automatically interpret the rest of inspection images based on supervised learning.

Many researches proposed some efficient approaches to extraction of textural features. Wunsch and Laine (1995) concluded wavelet descriptors insensitive to individual shape variations and better than Fourier descriptors in shape representation for handprinted characters. Amet, Ertüzün, and Erçil (2000) used wavelet transform (WT) and co-occurrence matrices to extract the co-occurrence features of defective textile fabrics which are powerful in detecting defects. Moreover, wavelet transform has multiresolution technique, so its sub-band decomposition of tree structure is appropriate for detection of local signal varieties on images. Arivazhagan and Ganesan (2003) presented that co-occurrence features computed from discrete wavelet transformed images are useful for texture segmentation. In this paper, the hybrid use of wavelet transform and co-occurrence matrices is considered as an effective solution for the texture analysis of sewage structural conditions. Based on the extracted co-occurrence features, the diagnostic system could be trained to assign each defect pattern to a correct category.

At present, pattern recognition techniques commonly use multilayer neural network and statistic methods such as Bayesian classifier, maximum likelihood method, and decision tree. Marchant and Onyango (2003) compared a Bayesian classifier with a multilayer feed-forward neural network to a plant/weed/soil discrimination case. The classification result demonstrated that the Bayesian classifier outperform the neural network due to its optimization in the sense of total misclassification error. Unfortunately, in most actual cases there is rarely a complete knowledge about the probabilistic structure of the problem. Thus, the nonparametric methods such as neural networks have been proposed to be widely applied to many actual pattern recognition problems (Duda, Hart, & Stork, 2001). Currently, back-propagation neural network (BPN), radial basis network (RBN), and support vector machine (SVM) are the three commonly used neural networks (Liao, Fang, & Nuttle, 2004) that is adopted in this paper to solve this classification problem of pipe defect patterns. The most efficient neural network technique is commended based on classification performance, and finally the application of this diagnostic system to a sewer system in the 9th district, Taichung City, Taiwan is discussed.

2. Methodology

2.1. Wavelet transform

Wavelet transform (WT) is a linear transform developed from Fourier transform. Unlike Fourier transform whose basis functions are sinusoids, wavelet transform is based on small waves, so-called wavelet, of varying frequency and limited duration so to obtain better resolutions along frequency scale (Chen, Wang, Yang, & McGreavy, 1999; Gonzalez & Woods, 2002). In multiresolution analysis (MRA), a scaling function is to create a series of approximation of a function or an image; additional functions, i.e. wavelet functions, are then used to encode the difference in information between adjacent approximations (Gonzalez & Woods, 2002). A set of wavelet functions is defined as:

$$\psi_{a,b}(x) = 2^{a/2}\psi(2^a x - b) \tag{1}$$

for all $a, b \in \mathbf{Z}$. $\mathbf{Z}$ is a set of integers. The scale parameter a controls stretch or compression of the mother wavelet function; the translation parameter b is an offset along the time axis; $2^{a/2}$ controls its height or amplitude (Chen et al., 1999; Gonzalez & Woods, 2002). Obviously a CCTV image can be regarded as the change of discrete signal along a two-dimensional (2D) scale. Hence, a 2D discrete WT (DWT) was proved to be useful for signal or image processing and pattern recognition (Bashar, Matsumoto, & Ohnishi, 2003; Hwang et al., 2005). The fast wavelet transform was considered as a computationally efficient implementation of the DWT was defined as (Gonzalez & Woods, 2002):

$$\varphi_{a,b}(x) = \sum_b h_\varphi(b)\varphi_{a+1,b}(x), \tag{2}$$

$$\psi_{a,b}(x) = \sum_b h_\psi(b)\varphi_{a+1,b}(x), \tag{3}$$

where $\varphi_{a,b}(x)$ and $\psi_{a,b}(x)$ are computed by convolving $\varphi_{a+1,b}(x)$ with the time-reversed scaling and wavelet vectors, $h_\varphi(b)$ and $h_\psi(b)$. In other words, the original function, $\varphi_{a+1,b}(x)$, is split into a lowpass (approximation compo-

nent) corresponding to $\varphi_{a,b}(x)$, and a highpass (detail component) corresponding to $\psi_{a,b}(x)$.

Due to the separability of wavelet transform, a 2D-DWT can be expressed in terms of a number of one-dimensional transforms (Parker, 1997). Through the decomposition of 2D-DWT, which is implemented by consecutive lowpass (L) and highpass (H) filtering through one-dimensional convolution, the CCTV image can be divided into an approximation image (LL) by a two-dimensional scaling function, $\varphi(x,y)$, shown as:

$$\varphi(x,y) = \varphi(x)\varphi(y) \tag{4}$$

and three detail images in horizontal (HL), vertical (LH), or diagonal (HH) directions by the three two-dimensional wavelets, $\psi^{H}(x,y)$, $\psi^{V}(x,y)$, and $\psi^{D}(x,y)$, respectively (see Fig. 1).

$$\psi^{H}(x,y) = \psi(x)\varphi(y), \tag{5}$$

$$\psi^{V}(x,y) = \varphi(x)\psi(y), \tag{6}$$

$$\psi^{D}(x,y) = \psi(x)\psi(y). \tag{7}$$

At each recurring step of decomposition, the approximation image is split into the next level of approximation and detail images until no more valuable information is obtained.

2.2. Co-occurrence matrices

A co-occurrence matrix is a square matrix in which each element M_{ij} in row (i) and column (j) directions records a relative occurrence frequency of a pair of pixels with the same gray level value separated by a certain pixel distance in one direction (Amet et al., 2000; Parker, 1997). For example, the approach of computing a co-occurrence matrix by one pixel distance in row direction is shown as Fig. 2. It is notable that the size of co-occurrence matrix depends on the range of the gray level values of the CCTV image. In common, there are 14 types of co-occurrence features derived from co-occurrence matrices useful for pattern classification (Haralick, Shanmugam, & Dinsten, 1973). However, the relationship between the co-occurrence features is not absolutely independent so that the redundant features would reduce the efficiency of pattern classification. Through a discriminant analysis, the most independent co-occurrence features with co-relation coefficients of less than 0.5 were found as entropy (ENT), correlation (COR), and cluster tendency (CLU), which were broadly used to describe the textures of objects on images (Amet et al., 2000; Arivazhagan & Ganesan, 2003; Hwang et al., 2005; Yao & Li, 2003), and are employed in this paper to describe the textures of sewer pipe defects and calculated as:

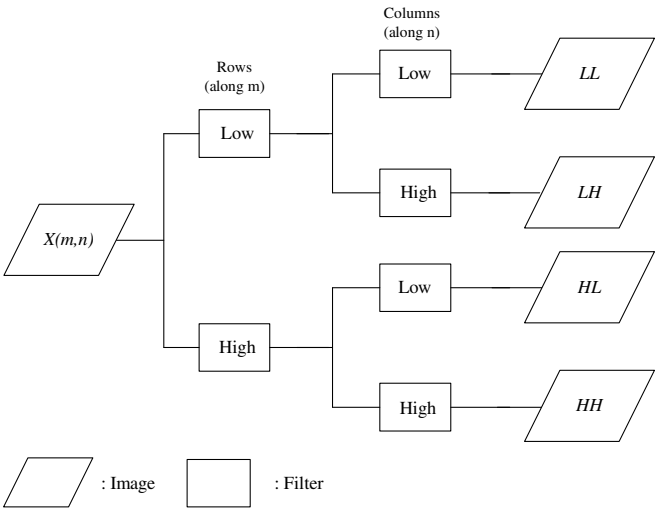


Fig. 1. Two-dimensional discrete wavelet transform.

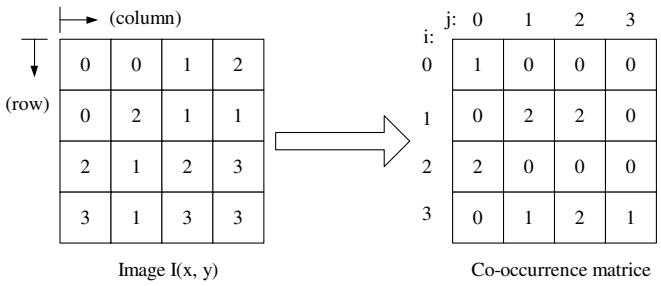


Fig. 2. Example of a horizontal co-occurrence matrix computation.

$$\mathrm{ENT} = -\sum_{i=1}^{n}\sum_{j=1}^{n} P_{ij}\cdot \log P_{ij}, \tag{8}$$

$$\mathrm{COR} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}(i\cdot j)P_{ij}, -\mu_x\mu_y}{\sigma_x\sigma_y}, \tag{9}$$

$$\mathrm{CLU} = \sum_{i=1}^{n}\sum_{j=1}^{n}(i-\mu_x+j-\mu_y)^2\cdot P_{ij}, \tag{10}$$

where

$$P_{ij} = \frac{M_{ij}}{\sum_{i=1}^{n}\sum_{j=1}^{n} M_{ij}} \tag{11}$$

$$\mu_x = \sum_{i=1}^{n}\sum_{j=1}^{n} i\cdot P_{ij} \tag{12}$$

$$\mu_y = \sum_{i=1}^{n}\sum_{j=1}^{n} j\cdot P_{ij} \tag{13}$$

$$\sigma_x = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}(i-\mu_x)^2\cdot P_{ij}} \tag{14}$$

$$\sigma_y = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}(j-\mu_y)^2\cdot P_{ij}}. \tag{15}$$

2.3. Back-propagation neural network (BPN)

BPN, a popular technique for pattern recognition and classification (Nolan, 2002), is one of the multilayer neural networks which usually consist of a three-layer structure:

input layer, hidden layer, and output layer. The numbers of neurons in the input and output layers are given according to the numbers of co-occurrence features and classes of pipe defects, respectively. As for the number of hidden layer neurons, there is no definite number and usually determined by trial and error (Moselhi & Shehab-Eldeen, 2000). Hidden layer neurons are fully connected to both input layer and output layer neurons, so the utility of hidden layer is to take in the external input or to send out the network output. Each hidden neuron computes its net activation based on the weighted sum of the input. Similarly, each output layer neuron computes its net activation based on the weighted sum of its net activations of hidden layer neurons. The type of the net activations between the input-hidden and hidden-output layers is defined as (Duda et al., 2001):

$$\text{net} = \sum_p X_p w_p + w_0 \tag{16}$$

where X_p represents either the co-occurrence features within the input-hidden layers or the activations of the input-hidden layers within the hidden-output layers; w_p is the interconnection weights; w_o is a bias; the subscript p indexes units in the layer. Each neuron within the hidden and output layers has a nonlinear sigmoid function of its activation shown as (Duda et al., 2001; Vapnik, 2000):

$$\begin{aligned}\text{Sgn}(\text{net}) = \tanh(\text{net}) &= \frac{\exp(\text{net}) - \exp(-\text{net})}{\exp(\text{net}) + \exp(-\text{net})} \\ &\equiv \begin{cases} 1 & \text{if} \quad \text{net} \geqslant 0 \\ -1 & \text{if} \quad \text{net} < 0 \end{cases}.\end{aligned} \tag{17}$$

During the training process, the optimal weights are estimated by minimizing the Least-mean-square (LMS) between the desired output t_l known a prior and the real output z_l. The LMS algorithm is adopted for evaluating the training error and expressed as

$$J(w) = \frac{1}{2}\sum_{l=1}^{c}(t_l - z_l)^2, \tag{18}$$

where c is the length of the output vectors; w represents all the weights in the network. Consequently, the basic approach of training a multilayer neural network is to input the training patterns of pipe defects to the input layer and to determine the assigned classes at the output layer via the network. The learning rule of BPN is to adjust the weights using the gradient descent of LMS for the sigmoid approximation of neural network. Initially, the weights are given randomly and adjusted in a direction by the chain rule for differentiation of $J(w)$ which involves two learning parameters, learning rate (η) and momentum factor (θ). Learning rate controls the relative size of the adjustment in weights. A large learning rate speeds up the rate of gradient descent, but decreases the stability of the network. Momentum factor is used to dynamically adapt the learning rate so that the danger of instability could be avoided and the convergence performance could be improved (Duda et al., 2001; Vapnik, 2000).

During BPN operation, the difference signal (t–z) is propagated backward through the network to adjust the weights using the gradient of LMS for the sigmoid approximation of neural network. The output vector has a dimension of $c \times n$, if n frames of CCTV images are inspected. Within an output vector, the lth element as "1" represents that this CCTV image is diagnosed as the lth defect pattern. Once the multilayer neural network is well trained through the above procedure, the intelligent diagnostic system of sewer pipes is ready for the recognition of pipe defect on a CCTV inspection image.

2.4. Radial basis network (RBN)

RBN, one of three-layer neural networks, is almost same as BPN, but with weights between the hidden and output layers. There are two stages, i.e. unsupervised learning and supervised learning, for training RBN (Zhang, Jiang, & Kamel, 2005). At the stage of unsupervised learning, clustering algorithm is used to divide all training samples of pipe defects s into subsets. The number of subsets is set as the number of neurons in the hidden layer of RBN. The characteristic of the training samples in each subset can be described by radial basis functions (G_j), one of activation functions, as (Hwang & Bang, 1997):

$$G_j = \begin{cases} \exp\left\{\frac{-\|s-c_j\|^2}{2\sigma_j^2}\right\}, & j = 1, 2, \ldots, n_{\text{H}}, \\ 1 & j = 0 \quad \text{(bais neuron)}, \end{cases} \tag{19}$$

where c_j and σ_j are the center and covariance matrix of each hidden neuron, respectively, which can be determined by clustering; $\|\cdot\|$ denotes the Euclidean distance; n_{H} represents the number of the hidden neurons. At the stage of supervised learning, a linear weighted sum between the hidden and output layers is computed as:

$$z_l = \sum_{j=0}^{n_H} G_j w_{lj}, \tag{20}$$

where l denotes the c classes of pipe defects; z_l is a linear weighted sum of the outputs of the hidden neurons; optimal w_{kl} is the solution of this set of linear equations by feed-forward calculation (Han & Xi, 2004; Liao et al., 2004; Zhang et al., 2005). A pipe defect pattern is fed into the trained RBN to be assigned into a certain class of pipe defect, so the output vector also has a dimension of $c \times n$, if n pipe defect patterns are recognized. Within an output vector, the lth element as a maximum z_l represents that this CCTV image is diagnosed as the lth pipe defect pattern.

2.5. Support vector machine (SVM)

SVM is a special type of feed-forward neural network. Given a set of training samples consisting of pairs of co-occurrence feature vector and class of pipe defect labels

as (X_i, Y_i), in which the subscript i indexes the lengths of the vectors ($Y_i \in \{1,-1\}$, $X_i \in \mathbf{R}^n$), the training approach of SVM is to adjust the weights ω and biases b ($b \in \mathbf{R}$) to search an optimal hyperplane and maximum margin which is defined as the distance of the closest vectors in both classes to the hyperplane. To find this hyperplane the following quadratic programming problem has to be solved (Scholkopf & Smola, 2002).

$$\text{Min} \quad \phi(\omega) = \frac{1}{2}(\omega \cdot \omega) = \frac{1}{2}\|\omega\|^2, \tag{21}$$

$$\text{s.t.} \quad Y_i[(X_i \cdot \omega) + b] \geqslant 1 \quad \text{for all } i = 1, \ldots, N. \tag{22}$$

However, it is difficult to straightforwardly solve the above primal model. Based on the duality theorem, the primal model can be transformed as the following dual model by modifying Eq. (21) with more convenience. To derive the dual problem, the Lagrangian is introduced,

$$\text{Max} \quad L(\omega, b, \lambda) = \frac{1}{2}\|\omega\|^2 - \sum_{i=1}^{N} \lambda_i \{Y_i[\langle X_i \cdot \omega \rangle + b] - 1\}, \tag{23}$$

where λ_i is Lagrange multipliers. A maximum L must satisfy the following conditions by applying a first-order derivation to Eq. (23) respective to b and ω as,

$$\frac{\partial L(\omega, b, \lambda)}{\partial b} = \sum_i \lambda_i^0 Y_i = 0, \quad \lambda_i^0 \geqslant 0, \tag{24}$$

$$\frac{\partial L(\omega, b, \lambda)}{\partial \omega} = \sum_i Y_i \lambda_i^0 X_i = \omega_0, \quad \lambda_i^0 \geqslant 0. \tag{25}$$

Only the Lagrange multipliers are non-zero at the saddle point and precisely meet the constraints of Eq. (22) as,

$$\lambda_i[Y_i(\langle X_i \cdot \omega \rangle + b) - 1] = 0 \quad \text{for all } i = 1, \ldots, N. \tag{26}$$

The patterns X_i for which $\lambda_i > 0$ are also called support vectors (SVs) which lie exactly on the margin (see Fig. 3), and all remaining training samples satisfy automatically their constraints Eq. (22). Substituting the conditions for Eqs. (24) and (25) into the Lagrangian Eq. (23), the following dual form is obtained as

$$\text{Max} \quad L_p(\omega, b, \lambda) = \sum_i \lambda_i - \frac{1}{2}\sum_{i,j} Y_i Y_j \lambda_i \lambda_j \langle X_i \cdot X_j \rangle. \tag{27}$$

In pattern recognition, a decision function, which correctly classifies the labeled samples (X_i, Y_i), is defined as (Scholkopf & Smola, 2002):

$$f_{\omega,b}(X) = \text{sgn}(\langle \omega \cdot X \rangle + b). \tag{28}$$

The way to find a large margin hyperplane for the linearly separable data by solving the above dual optimization problem is addressed above. However, to make SVM appropriate for real-world decision surfaces, several kernels are used to nonlinearly map the input data, X_i, from a sample space into a high-dimensional feature space where the implementation of linear separation becomes much easier (Cristianini & Shawe-Taylor, 2000; Vapnik, 2000). In this paper, well known kernel functions, radial Gaussian and polynomial, are used as (Brudzewski, Osowski, & Markiewicz, 2004; Scholkopf & Smola, 2002; Vapnik, 2000):

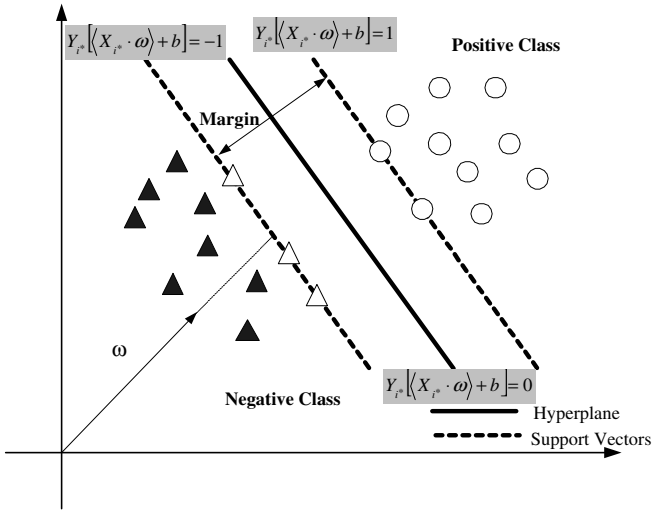


Fig. 3. Classification of two classes using SVM.

$$\text{Radial Gaussian}: \quad K(X, X_i) = \exp(-\|X - X_i\|^2/2\sigma^2), \tag{29}$$

$$\text{Polynomial}: \quad K(X, X_i) = (X^T \cdot X_i + \gamma)^d, \tag{30}$$

where K is kernel function; σ is the width of X_i; r is constant; d is the degree of polynomial.

A complete binary classification, where a class is labeled as either $+1$ or -1, has been discussed previously. Whereas, the pattern recognition of sewer pipe defects, one of many real-world problems, has more than two classes. Four methods including one versus the rest (also called one against all), pairwise classification, error-correcting output coding, and multi-class objective functions are able to deal with this issue (Scholkopf & Smola, 2002). One versus the rest, the most widely used by a set of binary c-class classifiers (Liao et al., 2004; Zheng, Li, & Song, 2004), repeats binary classifications to a multi-class object according to the maximal output (Scholkopf & Smola, 2002).

3. Study site and experimental materials

Water Research Centre (WRc) in UK categorizes major sewer pipe defects into 10 classes including open joint, displaced joint, crack, fracture, broken pipe, hole, collapse, spalling, wear, and deformation (Water Research Centre, 1993). This categorization rule has been widely applied to many sewage inspections and also to this study case. The recent inspection work was implemented in 2002 prior to a connection house program for the study site. Thousand frames of the CCTV inspection images were acquired and processed using 2D-DWT and computation of co-occurrence matrices described as follows.

3.1. Study site

The sewage in the 9th district of Taichung city, the largest city in the central Taiwan, consists of ten sewer systems, system A through J. The diameters of most pipelines are 300 mm and 85% of pipes are made of vitrified clay pipe (VCP). The other pipelines with diameters larger than 350 mm are made of reinforced concrete pipe (RCP). In 2002, the CCTV inspection work provides engineers 1101 CCTV inspection images for diagnosing the sewer pipe defects. The diagnosis result revealed that most of the detected pipe defects existed within system G so that in this paper it was selected to be the study site for demonstration. In system G, there were 291 CCTV inspection images acquired. A statistics for the 291 CCTV images reveals that open joint, crack, broken pipe, and fracture are the typical sewer pipe defects, and appears on CCTV images 107, 112, 16, and 56 times, respectively.

3.2. Experimental materials

The techniques of 2D-DWT and co-occurrence matrices were used to extract the textural features of the pipe defects. A 2-level 2D-DWT was used to divide each of

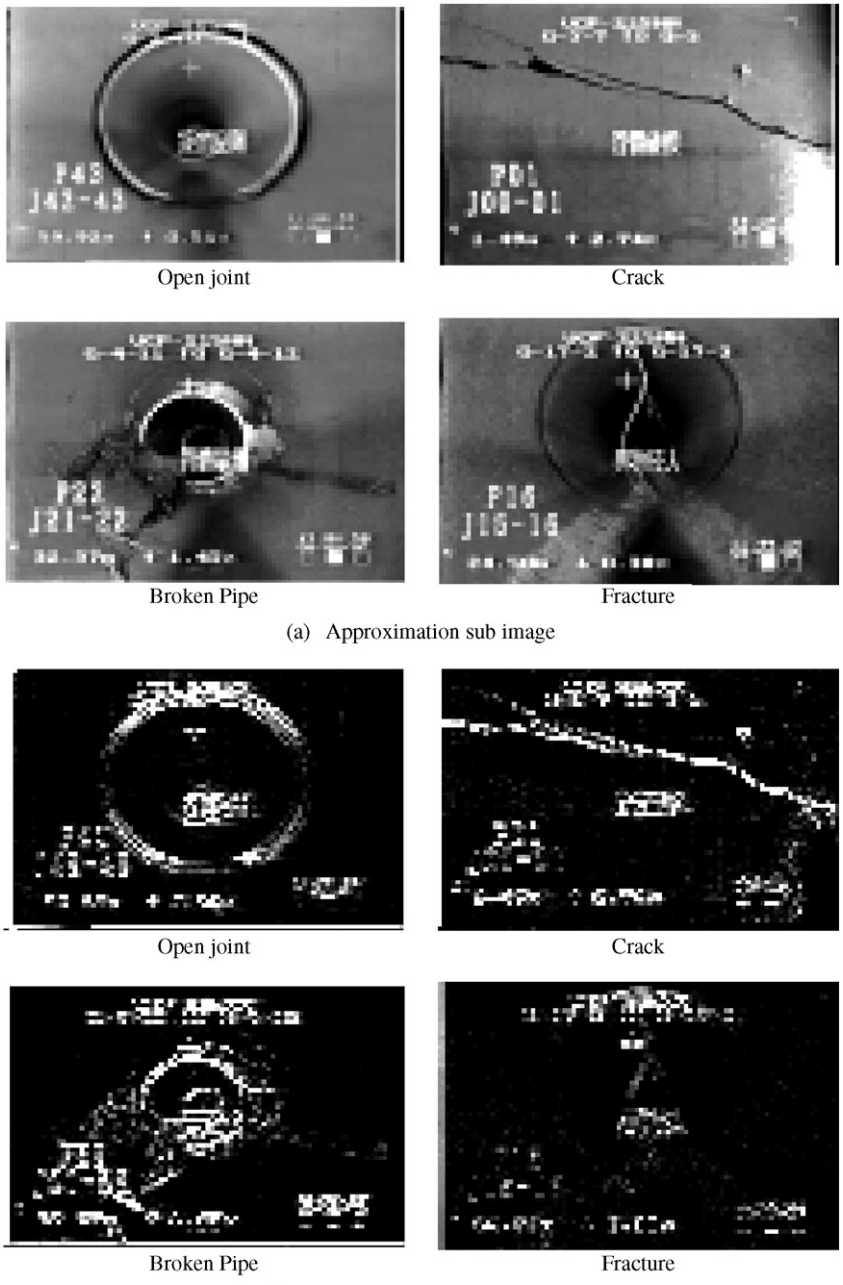
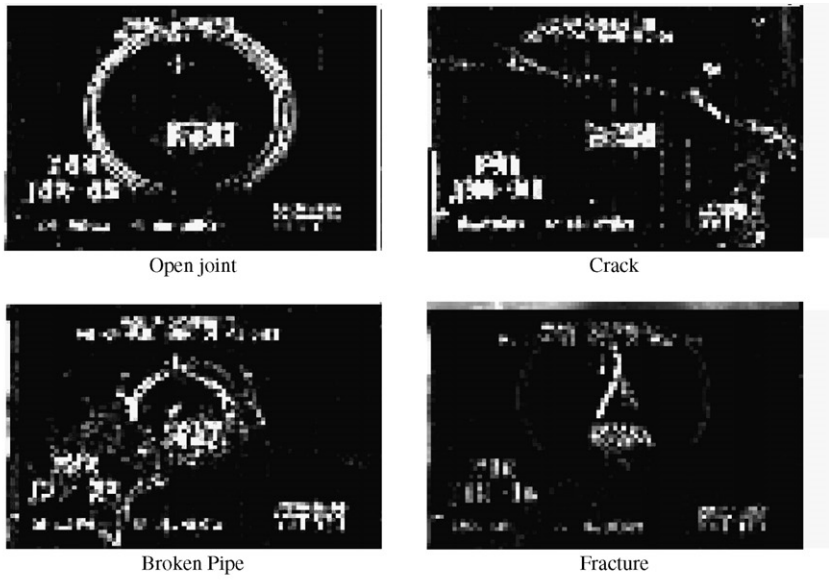
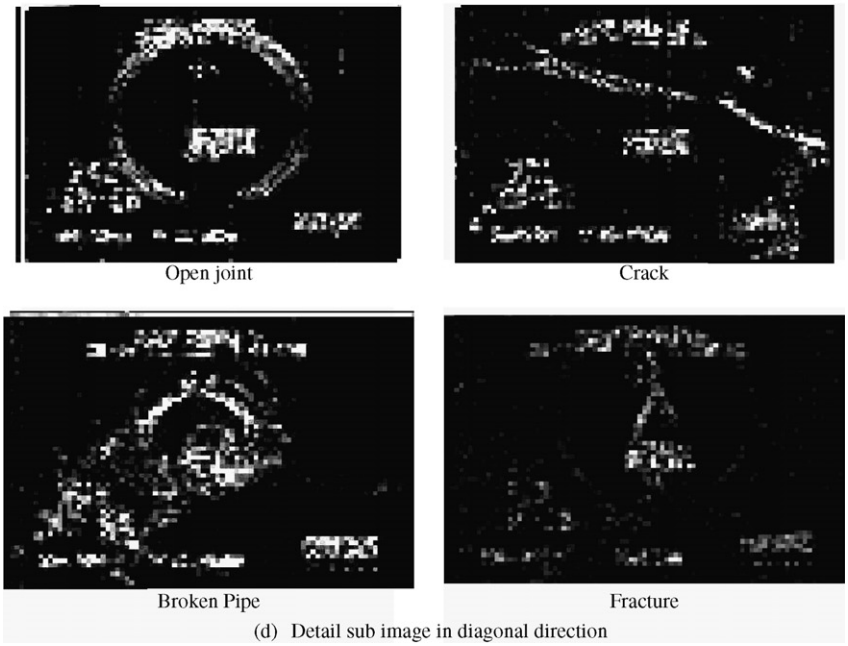


Fig. 4. 2-level 2-D wavelet transformation of the inspection images. (a) Approximation sub image (b) detail sub image in horizontal direction (c) detail sub image in vertical direction (d) detail sub image in diagonal direction.



(c) Detail sub image in vertical direction



(d) Detail sub image in diagonal direction

Fig. 4 (*continued*)

the 291 inspection images into four detail images including approximation sub image, horizontal detail sub image, vertical detail sub image, and diagonal detail sub image (see Fig. 4). Generally, distances of one pixel in the directions of 0°, 90°, and 45° and 135° are used to extract the co-occurrence features from the horizontal, vertical, and diagonal detail sub-images, respectively (Lohmann, 1995). In other words, the co-occurrence features in four orientations can be extracted from a CCTV image for the pipe defect. The co-occurrence feature in each orientation consists of entropy (ENT), correlation (COR), and cluster tendency (CLU). Thus, the average values of ENT, COR, and CLU in the four orientations were computed, respectively, to represent the textural feature of pipe defect. The data scatters of ENT vs. COR, ENT vs. CLU, and COR vs. CLU for the four pipe defects are shown as Figs. 5–7, respectively. These figures show that the extracted textural features have a large overlap among the four pipe defects. Thus, this pattern recognition problem is also a linearly non-separable problem.

Generally, training samples are either randomly selected among all patterns if all prior probabilities are equal or uniformly selected for each class if all prior probabilities are unknown so to avoid any selection bias (Brudzewski et al., 2004; Shehab & Moselhi, 2005; Shin, Lee, & Kim, 2005). Statistically, the number distribution of acquired

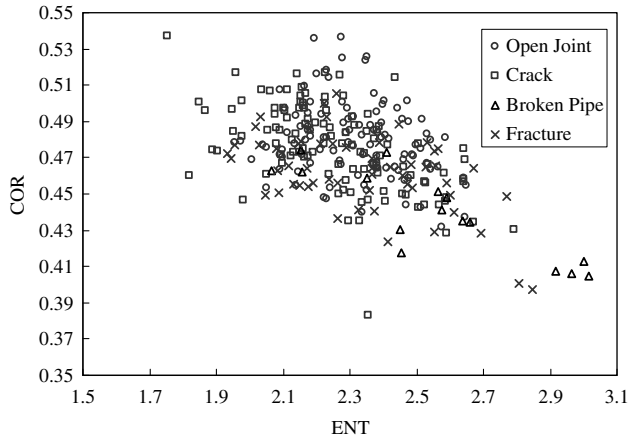


Fig. 5. Textural features of ENT vs. COR.

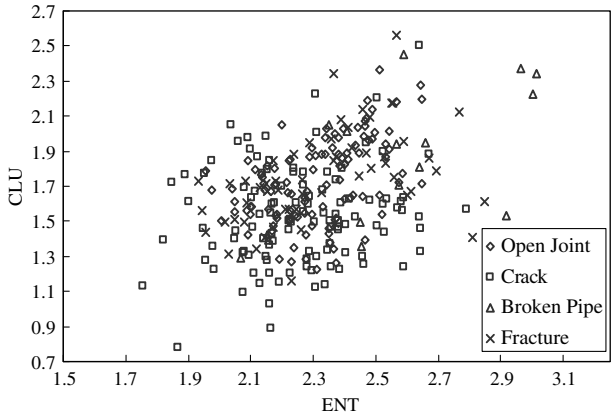


Fig. 6. Textural features of ENT vs. CLU.

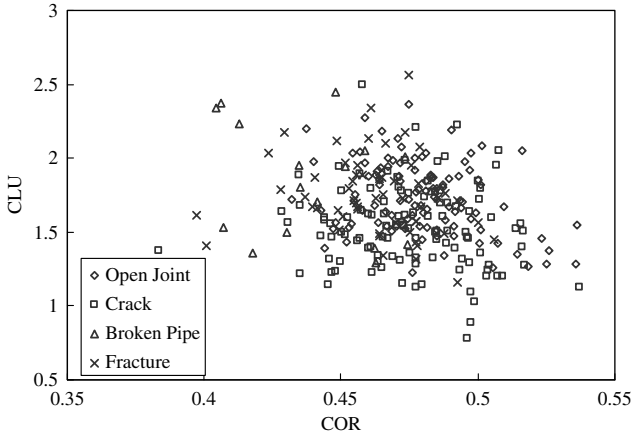


Fig. 7. Textural features of COR vs. CLU.

CCTV images in four pipe defects has a great difference. In the G sewer system, the pipe defect of broken pipe appears on 16 CCTV images only, which is comparatively less than the other pipe defects. Thus, two experiments (I and II) were designed to select 40 training images for the diagnostic system. In Experiment I, the training sample of each pipe defect pattern was assigned 10 images by assuming equal prior probability. Assuming unequal prior probability, Experiment II selected the training samples (also 40 images) of each pipe defect pattern based on its group size among the 291 images. Regarding to the sewage in the 9th district, the proportions of open joint, crack, broken pipe, and fracture are 36.8%, 38.5%, 5.5%, and 19.2%, respectively. Consequently, in Experiment II the numbers of training images of open joint, crack, broken pipe, and fracture were given by 15, 15, 2, and 8, respectively.

4. Implementation results and discussion

The concept of the diagnostic system proposed in this paper is shown as Fig. 8. The diagnosis performances of the three types of neural networks, BPN, RBN, and SVM, on pipe defects were compared with a discriminant analysis. Due to an expertise-based interpretation on the 291 CCTV images, four defect patterns can be discriminated by Bayesian classifier. BPN was established based on a supervised and LMS- correcting algorithm. The learning rate and momentum factor were set as $\eta = 0.5$ and $\theta = 0.999^t$ in which t represents the number of iterations ($t = 10{,}000$ in this paper to dynamically adapt η to 0 in the final iteration). According to trial-and-error results, the hidden layer of BPN was given by three neurons. The computation took about 450 s on Pentium IV 1.3G PC to obtain the classification result shown in Table 1. Compared with the classification result of Experiments I, the Experiment II has an overall accuracy of 54.6%, which is better than that (47.4%) of Experiment I. In Experiment II, the classification accuracy of the diagnostic system for pipe defects seems to vary with the number of training samples. The classification accuracy of broken pipe is zero that could result from a poor training due to comparatively less occurrence to other sewage defects.

RBN took about 0.1 s on Pentium IV 1.3G PC for the classification. Table 2 shows that the overall accuracy of Experiment II is also better than Experiment I. In Experiment II, the classification accuracy of each pipe defect seems to have a positive correlation with its proportion among the 291 CCTV images. Comparing Tables 1 and 2, there are three important results found from the experiments. Firstly, regarding to classification efficiency RBN performed much better than BPN due to taking extremely

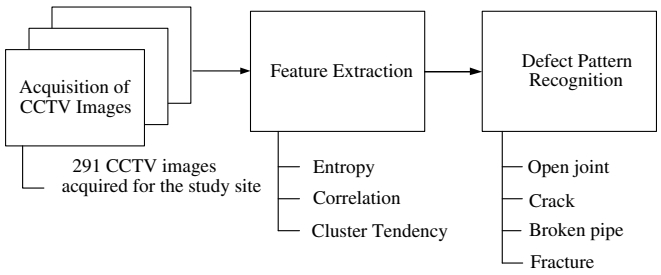


Fig. 8. Basic structure of automated classification.

Table 1
Classification accuracy of sewer pipe defects using the BPN classifier

Pipe defects	Experiment I (%)	Experiment II (%)
Open joint	35.5	66.4
Crack	24.1	67.8
Broken pipe	43.8	0.0
Fracture	67.9	21.4
Overall	47.4	54.6

Table 2
Classification accuracy of sewer pipe defects using the RBN classifier

Pipe defects	Experiment I (%)	Experiment II (%)
Open joint	30.0	68.2
Crack	35.7	57.1
Broken pipe	68.8	25.0
Fracture	50.0	28.6
Overall	38.1	54.0

short computation time. Secondly, RBN offered a better classification accuracy of 25% than BPN in the class of broken pipe in Experiment II. Thirdly, the overall accuracies in Experiment II are higher than Experiment I. In conclusion, RBN in Experiment II can perform the best classification.

SVM, a new classification method developed by Vapnik in 1999, finds an optimal hyperplane, which is orientated in such way that the margin is maximized, between any two classes (Liao et al., 2004; Vapnik, 2000). A less number of support vectors, usually about 10–30% of the total number of training samples, is expected so that the trained SVM could have a better generalization ability (Tran, Zhang, & Li, 2003). Originally SVM was designed for solving the problem of binary classification only (Liao et al., 2004), but recently has been modified for multi-classification, such as the problem of pipe defect diagnosis in this paper, by applying a "one against all" method (Liao et al., 2004).

Suitable for a linearly non-separable problem, SVM with either the radial Gaussian kernel or the polynomial kernel are applied to the automated classification of defect patterns. The parameters of the kernels affect the classification accuracies (Vapnik, 2000), and are determined based on heuristics during SVM training (Seo, 2007; Widodo, Yang, & Han, 2007). SVM took about one second on Pentium IV 1.3G PC for Experiments I and II to obtain classifications shown as Tables 3 and 4, respectively. Comparing Tables 3 with 4, SVM classification with different kernel functions obtained similar overall accuracies, whereas overall accuracies appear quite different and depend on the selection of training samples. Moreover, these testing results demonstrate that the obtained classification accuracy of each pipe defect in Experiment II is less affected by the selection of kernel functions than that in Experiment I.

The acquired numbers of support vectors in Experiments I and II are given in Tables 5 and 6, in which the

Table 3
Classification accuracy of sewer pipe defects using the SVM classifier in Experiment I

Pipe defects	Polynomial $(d, r) = (5, 3.3)$ (%)	Radial Gaussian $\sigma = 0.5$ (%)
Open joint	46.7	11.2
Crack	40.2	90.2
Broken pipe	87.5	50.0
Fracture	48.2	28.6
Overall	46.7	47.1

Table 4
Classification accuracy of sewer pipe defects using the SVM classifier in Experiment II

Pipe defects	Polynomial $(d, r) = (3, 1.2)$ (%)	Radial Gaussian $\sigma = 0.3$ (%)
Open joint	73.8	76.6
Crack	62.5	58.0
Broken pipe	18.8	18.8
Fracture	35.7	23.2
Overall	60.0	56.0

ratios of the total support vectors to samples are either within the acceptable range of 10–30% or approximately 10%. Basically SVM in Experiments I or II could be trained to be with superior generalization ability. With SV ratios over 50.0% (see Table 5), the defect pattern of broken pipe classified by SVM in Experiment I was over fitted due to relative less samples. Comparing the classification performances of RBN and SVM in Experiment II (see Tables 2 and 4), it can be found that the overall accuracies of SVM are little better than RBN. As for the efficiencies of classification, RBN or SVM both took an extremely short

Table 5
Numbers of support vectors (SV) in Experiment I

Pipe defects	Polynomial $(d, r) = (5, 3.3)$		Radial Gaussian $\sigma = 0.5$	
	No. of SV	% of SV	No. of SV	% of SV
Open joint	12	11.2	7	6.5
Crack	16	14.3	3	2.7
Broken pipe	14	87.5	8	50.0
Fracture	14	25.0	8	14.3
Total	56	19.2	26	8.9

Table 6
Numbers of support vectors (SV) in Experiment II

Pipe defects	Polynomial $(d, r) = (3, 1.2)$		Radial Gaussian $\sigma = 0.3$	
	No. of SV	% of SV	No. of SV	% of SV
Open joint	9	8.4	7	6.5
Crack	10	8.9	9	8.0
Broken pipe	6	37.5	4	25.0
Fracture	7	12.5	7	12.5
Total	32	11.0	27	9.3

Table 7
Classification accuracy of sewer pipe defects using Bayesian classifier with [ENT, COR, CLU]

		Classification				Column total
		Open joint	Crack	Broken pipe	Fracture	
Reference data	Open joint	77	26	2	2	107
	Crack	34	72	3	3	112
	Broken pipe	2	5	6	3	16
	Fracture	27	14	3	12	56
Row total		140	117	14	20	291
Accuracy (%)		72.0	64.3	37.5	21.4	57.4

time for computation. In conclusion, RBN or SVM in Experiment II could obtain better classification results for the pattern recognition of sewer pipe defects.

In order to verify the classification performance of the diagnostic system, a discriminant analysis based on Bayesian theory was done for the 291 sewer pipe images through the feature vector [ENT, COR, CLU]. The analysis result shown as Table 7 shows that the overall accuracy of 57.4% is extremely close to the overall accuracies obtained by the SVM. The overall accuracy of 60.0% obtained by the SVM with the Polynomial kernel is even better than the discriminant analysis result.

5. Conclusions

Neural networks, BPN, RBN, and SVM, have been applied to many classification problems, but only BPN ever was applied to the detection and classification of pipe defects. This paper is the first trial to apply RBN and SVM to the diagnosis of sewer pipe defects. As input elements to BPN, RBN, and SVM, the textural feature vectors consisting of the average ENT, COR, and CLU extracted by 2D-DWT and co-occurrence matrices were adopted to represent the texture information of sewer pipe defects on the CCTV images. The larger the size of co-occurrence matrix is, the more computation time takes. In this paper, 291 frames of CCTV inspection images of sewer system G in the 9th district of Taichung City were used to implement the automated classification of sewage structural conditions. On Pentium IV 1.3G PC, the testing result shows that about 60 s are required for a defect pattern to obtain co-occurrence features.

Two experiments were designed to verify the diagnostic system due to the great difference of the acquired images numbers between different pipe defect patterns. The testing results reveal that Experiment II, in which the training sample number of each pipe defect pattern was decided based on its group size, is more suitable for the diagnostic system. Based on these experiments, an automated diagnosis of sewer pipe defects should be firstly implemented an expertise-based human diagnosis to a portion of all the CCTV images (about 1/7 of all images or 40 images suggested in this paper) for the neural network training. Subsequently, the diagnose system can be applied to the rest of CCTV images. In addition, it is found that the performance of the diagnostic system using either SVM or RBN is better than BPN. Especially for SVM, however, the best parameters within the kernel functions are usually determined based on heuristics at present. Some literatures suggested that a structured method of selecting the best learning parameters in SVM should be developed for the best classification performance (Seo, 2007).

To compare the best accuracy (60%) resulted form SVM, a Bayesian classifier based on discriminant analysis was also used to diagnose the 291 pipe defect images and obtained an accuracy of 57.4% that proves the utility of the diagnostic system. Some textures of inspection attributes manually recorded on the original CCTV images (see Fig. 4) become harmful noises so to derive the extracted textural features of pipe defects with a linear non-separable relationship for the automated diagnosis and to increase the difficulty of classification. If the textural features of pipe defects are extracted from pure CCTV images without inspection records attached, the diagnosis accuracy should be increased. However, the accuracy of the automated diagnosis for sewer pipe defects remains to be improved for practical application in the future work.

Acknowledgement

This study in part was sponsored by the Construction Bureau, Construction and Planning Agency, Ministry of the Interior, Taiwan. This author would like to thanks Pipewell Ltd., Taiwan for providing the CCTV inspection images.

References

Amet, A. L., Ertüzün, A., & Erçil, A. (2000). An efficient method for texture defect detection: sub-band domain co-occurrence matrices. *Image and Vision Computing, 18*(6–7), 543–553.

Arivazhagan, S., & Ganesan, L. (2003). Texture segmentation using wavelet transform. *Pattern Recognition Letters, 24*(16), 3197–3203.

Bashar, M. K., Matsumoto, T., & Ohnishi, N. (2003). Wavelet transform-based locally orderless images for texture segmentation. *Pattern Recognition Letter, 24*(15), 2633–2650.

Brudzewski, K., Osowski, S., & Markiewicz, T. (2004). Classification of milk by means of an electronic nose and SVM neural network. *Sensors and Actuators B: Chemical, 98*(2–3), 291–298.

Chen, B. H., Wang, X. Z., Yang, S. H., & McGreavy, C. (1999). Application of wavelets and neural networks to diagnostic system development, 1, feature extraction. *Computers and Chemical Engineering, 23*(7), 899–906.

Cordes, S., Berns, K., Eberl, M., Ilg, W., & Suna, R. (1997). Autonomous sewer inspection with a wheeled, multiarticulated robot. *Robotics and Autonomous Systems, 21*(1), 123–135.

Cristianini, N., & Shawe-Taylor, J. (2000). *An introduction to support vector machines and other kernel-based learning methods*. Cambridge, NY: Cambridge University Press.

Duda, R. O., Hart, P. E., & Stork, D. G. (2001). *Pattern classification*. New York: Wiley.

Fenner, R. A. (2000). Approaches to sewer maintenance: a review. *Urban Water, 2*(4), 343–356.

Gonzalez, R. C., & Woods, R. E. (2002). *Digital image processing*. Englewood Cliffs, NJ: Prentice-Hall.

Han, M., & Xi, J. (2004). Efficient clustering of radial basis perceptron neural network for pattern recognition. *Pattern Recognition, 37*(10), 2059–2067.

Haralick, R. M., Shanmugam, K., & Dinsten, I. (1973). Texture features for image classification. *IEEE Transactions on Systems, Man and Cybernetics, 3*(6), 610–621.

Hwang, Y. S., & Bang, S. Y. (1997). Recognition of unconstrained handwritten numerals by a radial basis function neural network classifier. *Pattern Recognition Letters, 18*(7), 657–664.

Hwang, H. G., Choi, H. J., Kang, B. D., Yoon, H. K., Kim, H. C., Kim, S. K., & Choi, H. K. (2005). Classification of breast tissue images based on wavelet transform using discriminant analysis, neural network and SVM. In Proceedings of the 7th international workshop on enterprise networking and computing in healthcare industry (pp. 345–349). Busan, Korea.

Liao, Y., Fang, S. C., & Nuttle, H. L. W. (2004). A neural network model with bounded-weights for pattern classification. *Computers and Operations Research, 31*(9), 1411–1426.

Lohmann, G. (1995). Analysis and synthesis of textures: A co-occurrence-based approach. *Computers & Graphics, 19*(1), 29–36.

Madryas, C., & Przybyla, B. (1998). Inspection of pipes as an element of operating municipal sewerage networks. *Tunneling and Underground Space Technology, 13*(1), 57–64.

Makar, J. M. (1999). Diagnostic techniques for sewer systems. *Journal of Infrastructure Systems, 5*(2), 69–78.

Marchant, J. A., & Onyango, C. M. (2003). Comparison of a Bayesian classifier with a multilayer feed-forward neural network using the example of plant/weed/soil discrimination. *Computers and Electronics in Agriculture, 39*(1), 3–22.

McKim, R. A., & Sinha, S. K. (1999). Condition assessment of underground sewer pipes using a modified digital image processing paradigm. *Tunneling and Underground Space Technology, 14*(2), 29–37.

Moselhi, O., & Shehab-Eldeen, T. (1999). Automated detection of surface defects in water and sewer pipes. *Automation in Construction, 8*(5), 581–588.

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

Nolan, J. R. (2002). Computer systems that learn: An empirical study of the effect of noise on the performance of three classification methods. *Expert Systems with Applications, 23*(1), 39–47.

Parker, J. R. (1997). *Algorithms for image processing and computer vision*. New York: Wiley.

Scholkopf, B., & Smola, A. J. (2002). *Learning with kernels: Support vector machines, regularization, optimization, and beyond*. Cambridge, MA: MIT Press.

Seo, K. K. (2007). An application of one-class support vector machines in content-based image retrieval. *Expert Systems with Applications, 33*(2), 491–498.

Shehab, T., & Moselhi, O. (2005). Automated detection and classification of infiltration in sewer pipes. *Journal of Infrastructure Systems, 11*(3), 165–171.

Shin, K. S., Lee, T. S., & Kim, H. J. (2005). An application of support vector machines in bankruptcy prediction model. *Expert Systems with Applications, 28*(1), 127–135.

Sinha, S. K., & Fieguth, P. W. (2006). Segmentation of buried concrete pipe images. *Automation in Construction, 15*(1), 47–57.

Tran, Q. A., Zhang, Q. L., & Li, X. (2003). Reduce the number of support vectors by using clustering techniques. In *Proceedings of the 2nd international conference on machine learning and cybernetics* (pp. 1245–1248). Xi'an, China.

Vapnik, V. N. (1999). *The natural of statistical learning theory*. New York: Springer.

Vapnik, V. N. (2000). *The nature of statistical learning theory*. New York: Springer.

Water Research Centre (1993). *Sewer rehabilitation manual*. Swindon, UK: Water Research Centre/Water Authorities Association.

Widodo, A., Yang, B. S., & Han, T. (2007). Combination of independent component analysis and support vector machines for intelligent faults diagnosis of induction motors. *Expert Systems with Applications, 32*(2), 299–312.

Wirahadikusumah, R., Abraham, D. M., Iseley, T., & Prasanth, R. K. (1998). Assessment technologies for sewer system rehabilitation. *Automation in Construction, 7*(4), 259–270.

Wunsch, P., & Laine, A. F. (1995). Wavelet descriptors for multiresolution recognition of handprinted characters. *Pattern Recognition, 28*(8), 1237–1249.

Xu, K., Luxmoore, A. R., & Davies, T. (1998). Sewer pipe deformation assessment by image analysis of video surveys. *Pattern Recognition, 31*(2), 169–180.

Yang, M. D., & Su, T. C. (2007). An optimization model of sewerage rehabilitation. *Journal of the Chinese Institute of Engineers, 30*(4), 651–659.

Yang, M. D., & Su, T. C. (2006). An automation model of sewerage rehabilitation planning. *Water Science and Technology, 54*(11–12), 225–232.

Yao, H., & Li, B. (2003). An efficient approach for texture-based image retrieval. In *Proceedings of 2003 IEEE international workshop on neural networks for signal* (pp. 1039–1043). Toulouse, France.

Zhang, C., Jiang, J., & Kamel, M. (2005). Intrusion detection using hierarchical neural networks. *Pattern Recognition Letters, 26*(6), 779–791.

Zheng, E., Li, P., & Song, Z. (2004). Performance analysis and comparison of neural networks and support vector machines classifier. In *Proceedings of the 5th world congress on intelligent control* (pp. 4232–4235). Hangzhou, China.