https://dx.doi.org/10.1016/j.autcon.2016.08.033

A Self Organizing Map Optimization based Image Recognition and Processing Model for Bridge Crack Inspection

Jieh-Haur Chen[1], Mu-Chun Su[2], Ruijun Cao[3], and Shu-Chien Hsu[4,*] Jin-Chun Lu[5]

The current deterioration inspection method for bridges heavily depends on human recognition, which is time consuming and subjective. This research adopts Self Organizing Map Optimization (SOMO) integrated with image processing techniques to develop a crack recognition model for bridge inspection. Bridge crack data from 216 images was collected from the database of the Taiwan Bridge Management System (TBMS), which provides detailed information on the condition of bridges. This study selected 40 out of 216 images to be used as training and testing datasets. A case study on the developed model implementation is also conducted in the severely damage Hsichou Bridge in Taiwan. The recognition results achieved high accuracy rates of 89% for crack recognition and 91% for non-crack recognition. This model demonstrates the feasibility of accurate computerized recognition for crack inspection in bridge management.

Keywords: Self Organizing Map Optimization, Image recognition, Bridge inspection

[1] Distinguished Professor, Institute of Construction Engineering and Management, National Central University, Taiwan.

[2] Professor, Department of Computer Science and Information Engineering, National Central University, Taiwan.

[3] Ph.D. Student, Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hong Kong.

[4] Assistant Professor, Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hong Kong. *Corresponding Author.

[5] Graduate Student, Institute of Construction Engineering and Management, National Central University, Taiwan.

1. Introduction

In 2009, Typhoon Morakot wrought catastrophic damage on Taiwan, leaving 461 people dead and 192 others missing, with a cost of roughly 110 billion New Taiwan dollars (NTD), which is close to 3.3 billion United States dollars (USD) in damages. The extreme amount of rain triggered enormous mudslides and flooding throughout southern Taiwan. Typhoon Morakot not only tested how the Taiwan government could relieve the victims of a severe disaster, but also drew attention to the need to improve the safety of infrastructure to reduce the impact of disasters. There are over twenty thousand bridges located across Taiwan. As bridges have an important role in facilitating transportation, damage to bridges by disasters not only threatens the safety of users, but can also disrupt traffic flows and cause residents to be locked in place.

Cracking in concrete bridges is an inevitable problem resulting from natural processes and can invite spectacular failure of the entire bridge. Cracks not only provide access to harmful and corrosive chemicals inside concrete, but also allow water and deicing salts to penetrate through bridge decks, which can damage superstructures and bridge esthetics. Routine inspections are widely adopted and are carried out manually by certified bridge inspectors every two years, as stipulated in the National Bridge Inspection Standard by Federal Highway Administration (FHWA) in USA. Inspection results are mainly based on the inspectors' observations and visual assessment [1, 2]. However, such bridge detection methods have several limitations. The inspection process is laborious, time-consuming, and influenced by the subjective behavior of individual inspectors. The visual inspection only provides qualitative information on defects. Moreover, finding experienced bridge inspectors poses a challenge for the construction industry, which is now facing a pressing shortage of experienced and highly trained inspection personnel [2, 3].

In order to overcome these issues, considerable research has been conducted in an effort to develop automated bridge crack inspection tools to reduce the field work required for inspectors [4]. For example, Oh and Jang et al. (2009) suggest certain image processing algorithms for detecting and tracing cracks combined with the use of a robot mechanism [5]. Zhu et al. (2010) propose an automated bridge condition assessment system with a focus on detecting large-scale bridge concrete columns [6]. Yu et al. (2012) designed a robot which can detect fissures underneath a bridge. They provided a safe and effective machine vision technology to detect the bridge [7]. Bu et al. (2014) developed an automatic bridge inspection approach by employing Support Vector Machines to classify cracks based on wavelet-based image features. The researchers tested 50 different image samples, and both 'complex' and 'normal' images were considered. The resulting recognition accuracy rates of the crack ranged from 74% to 93.26%, varying according to the different image types, training set types, and the feature extraction methods used [8]. Li, et al. (2014) also put forward a method consisting of crack extraction, an electronic distance measurement algorithm, and an image segmentation algorithm to detect cracks [3]. Further increasing the reliability and accuracy of the results remains an ongoing effort in this research area. The algorithm and the restricted conditions employed in this study contribute to this effort to improve crack detection. The objective of this study is to develop an automatic bridge crack detection method based on self-organizing map optimization through image processing technology. Several stages of image recognition for bridge cracks are utilized to process 216 images of randomly selected bridges located in northern Taiwan. The model presented here can improve the accuracy of bridge crack inspections and reduce the cost of labor.

2. Bridge Crack Inspection

Degradation often occurs during the final stage of reinforcement concrete structures' life cycle. The various degrees of maintenance and service conditions of structures in different natural environments derive from disparate degradation rates and consequences. In Taiwan, bridges very easily deteriorate due to high humidity, frequent earthquake loading, and overloading by heavy vehicles [9]. The acceleration of degradation in reinforced concrete structures can be attributed to natural factors and human factors. The natural factors causing cracks in and damage to the structures mainly include strong winds, storm erosion, earthquakes, flooding, and other force majeure, while human factors may be the improper or wrong use of the structures, such as vehicle overload and irregular maintenance. The cracks on surfaces often show the initial fatigue reaction of the bridge components and forecast the failure of reinforced concrete structures [10]. Once the crack emerges, the degradation of the entire structure will follow in a short time, reducing in the overall strength and durability of the structure. Thus, in order to extend the service life of concrete structures, timely monitoring and remedial management are extremely necessary to avoid more serious deterioration.

Current bridge crack inspection systems, implementation methods, and rating records for detection vary among countries all over the world. The Federal Highway Administration (FHWA) specifies four condition states to evaluate the elements of a bridge: Good, Fair, Poor, and Severe [1], and also employs a scale of 0-9 and NA to rate the National Bridge Inventory (NBI) conditions [2]. The Japan Highway Public Corporation classifies the performance degradation of bridge elements as I, II, III and IV [11]. However, the degradation level should be considered first before rating. For instance, when a crack exists in components, the crack shape, amplitude, and interval are key aspects that should be evaluated to assign more accurate ratings. A standardized bridge inspection system has not yet been formulated in Taiwan. However, local bridge authorities have developed their distinctive and practical assessment criteria and rating record modes.

In Taiwan, the predominant nondestructive evaluation method currently employed is visual inspection conducted by professional inspectors, with its relative advantages in cost and speed. The DERU evaluating method is a visual inspection assessment approach to bridge management developed by the Taiwan Area National Freeway Bureau, which divides component degradation into the degree of degradation (Degree), the scope of degradation (Extend), the importance of the degradation phenomenon to components (Relevancy) and maintenance urgency (Urgency) four parts with employing 4 levels (shown in Table 1) to evaluate [9]. The DERU criterion enables bridges to be evaluated in as short a time as possible, which greatly enhances inspection efficiency. However, visual inspection to a great extent relies on the naked-eye observation of the component appearance to judge the degradation degree and scope. Or in other words, this method completely depends on the subjective evaluation of the inspector. Considering the number of bridges in total, the increasing number of bridges damaged by natural disasters in Taiwan in particular, the manpower shortage affecting related bridge management institutions, and the difficulty in training learners with relevant professional knowledge, carrying out regular, effective bridge inspections has become an enormous challenge. Especially after serious natural disasters, the workload of inspectors is even more onerous. Thus, the results of visual inspection are quite biased due to the different inspecting habits of individuals, and often have questionable reliability.

Table 1 DERU Rating System criteria

	0	1	2	3	4

D	Not applicable	Good	Fair	Poor	Severe
E	Not applicable	< 10%	< 30%	< 60%	<
R	Not applicable	Marginal	Small	Medium	Large
U	Not applicable	Routine Maintenance	3 years	1 year	Urgent Maintenance

3. Data Collection

This research employs a digital camera to capture concrete bridge cracks in order to develop an image recognition program and process image identification. Before shooting, the choice of the bridges included in the sample for the present study was based on the DERU visual inspection results. Ten concrete bridges in northern Taiwan are selected, and the crack images were randomly chosen from the artificial field shooting database. According to the manual of highway maintenance which was published by the Taiwan Highway Administration, routine highway maintenance needs to be conducted every two years [12]. Thus the bridges that have been maintained recently were filtered out from this research. According to the DERU bridge detection assessment criteria issued by the Join Engineering Consultants, when the D value is greater than 2 (Degree value refers to the severity degree of bridge degradation, from grade 1 to grade 4), the damage intensity will be visible in appearance and maintenance is needed.

The training samples used to research and develop the image recognition system of this study are the close-range images of concrete bridge cracks documented by the handheld digital cameras. Field environmental conditions must be taken into careful consideration when shooting onsite to obtain the images for training purposes. Since image processing is necessary before recognition, greater uniformity and consistency of the image acquisition conditions is desired to avoid errors in the image processing. The intensity of illumination in shooting is the major element influencing the digital camera photos. Changes in the sun position give rise to a variety of natural illuminations at each time point. Additionally, the intensities of illumination generated by natural light and artificial light are entirely different, specifically the illumination intensity of artificial light is more than that of natural light by several-fold.

Generally speaking, due to the light and shadow, the color of a concrete structure crack is deeper and darker compared to the color of the surrounding surface. Natural light and auxiliary light is critical in affecting the black shadow area of the cracks, as well as reflected light sources on the concrete surfaces, all impacting the system recognition results. Because of the uncontrollability of the climate, seasons, and time with respect to the nature light in the field, more easily controlled artificial lighting is employed to assist with shooting. In order to reduce the effects of natural light in this study, a natural light shield was applied to completely eliminate the uncontrolled factors of natural light. Setting the camera flash and fixing the shooting angle play a significant role in standardizing the artificial light source. Thus the taking lens was fixed to be perpendicular to the degradation structure plane in this study.

According to the user manual of the digital camera used in this study, the size of each color image is 3088×2056 pixels, which is more convenient for storage and clear enough for identifying the subjects. Excessively high-resolution photos may cause redundancy of image details in the training process. Thus, the sample images obtained for this study are all based on the above conditions. The guiding principle of the crack recognition program in this research is to facilitate image processing from the color difference between cracks and bridge surface for computer recognition. In order to put this program into practice, the classification rule must be set up through the training procedure. After screening the bridges with respect to the Taiwan Bridge Management System (TBMS), ten bridges were randomly selected as the training sample with 20-30 pictures for each bridge. For the training stage, special attention was paid to shoot on-site such that photos would only contain a main line crack while avoiding shots of multiple cracks existing at the same time. The 216 sequentially numbered image samples were acquired assuming the above constraints. From this set of 216 images a random sample of 40 images was selected, with 36 used as the training set and the other 4 comprising the recognition set. For the purpose of demonstrating that the program in this study can be effectively applied in practice, the case study portion of the current research used 18 images from the image samples of Hsichou Bridge in Taiwan and assessed by the recognition program.

4. Development of Image Recognition and Processing Model

The proposed image recognition and processing model was developed on the basis of the concepts of self-organizing feature map optimization (SOMO), fuzzy logic control, and hyper-rectangular composite neural networks (HRCNNs). The SOMO was developed by Su et al. in 2004 [13] and has been applied to several areas such as construction sequencing for building renovation and secant pile walls [14, 15]. The model development in this study starts with the creation of HRCNN integrated with fuzzy logic. The HRCNN is derived fundamentally from artificial neural networks (ANN) using the rule extraction concept to achieve machine learning and pattern classification. Adopting the supervised decision-direct learning (SDDL) algorithm, the HRCNN classifier can attain a 100% classification rate [16]. Assuming that the output is expressed as *Out(x)*, the output function is written as:

$$Out(\underline{x}) = f\left(\sum_{j=1}^{J} Out_j(\underline{x}) - \eta\right) \quad (1)$$

$$Out_j(\underline{x}) = f\left(net_j(\underline{x})\right) \quad (2)$$

$$net_j(\underline{x}) = \sum_{i=1}^{p} f\left((M_{ji} - x_i)(x_i - m_{ji})\right) - p \quad (3)$$

$$f(x) = \begin{cases} 1 & if \quad x \geq 0 \\ 0 & if \quad x < 0 \end{cases} \quad (4)$$

where Out(x) belongs to Rp $\rightarrow \{0, 1\}$; η is a small positive real number; x = (x1, . . . ,x_p)^T stands for an input pattern; M_{ji} and m_{ij} $\in$ R are the adjustable synaptic weights of the *j*th neuron of the hidden layer; *p* is the dimension of the input variable. Once $\underline{x}$ is in one of the *J* hyper-rectangular areas, *Out*($\underline{x}$) = 1; otherwise, *Out(x)* = 0. The values of the corresponding synaptic weights in the J hidden nodes of a trained HRCNN are interpreted as IF-THEN rules:

$$\begin{aligned} &\text{If } \left(\underline{x} \in [m_{11}, M_{11}] \times \cdots \times [m_{1p}, M_{1p}]\right) \text{ Then } Out(\underline{x}) = 1, \\ &\ldots \\ &\text{ELSE If } \left(\underline{x} \in [m_{J1}, M_{J1}] \times \cdots \times [m_{Jp}, M_{Jp}]\right) \text{ Then } Out(\underline{x}) = 1. \\ &\text{ELSE } Out(\underline{x}) = 0 \end{aligned} \quad (5)$$

With fuzzy logic added, the mechanism for HRCNN is subject to change where *mj*($\underline{x}$) is employed to replace Eq. (4) so as to achieve measurement of similarity between the inputs and the hyper-rectangular area. Hence, Eq. (4) is changed to:

$$m_j(\underline{x}) = \exp\left\{-s_j^2\left[per_j(\underline{x}) - per_j\right]^2\right\} \quad (6)$$

where

$$per_j = \sum_{i=1}^{p}\left(M_{ji} - m_{ji}\right) \quad (7)$$

$$per_j(\underline{x}) = \sum_{i=1}^{p} \max\left(M_{ji} - m_{ji}, x_i - m_{ji}, M_{ji} - x_i\right) \quad (8)$$

The output is re-written as:

$$Out(\underline{x}) = \sum_{j=1}^{J} w_j m_j(\underline{x}) + \theta \quad (9)$$

where w_j is the weight of the jth neuron of the hidden layer; s_j is the sensitivity; and θ is an adjustable value. The fuzzy based HRCNN (FHRCNN is capable of yielding linearly weighted rules when compared to Eqs. (2) and (9). That is because $mj(\underline{x})$ is more flexible and can be either Gaussian or a Step function. Therefore, Eq. (5) is subject to modification by:

$$\begin{array}{ll} \text{If } \left(\underline{x} \text{ is } HR_1\right) & \text{Then } Out(\underline{x}) \text{ is } w_1 \\ \ldots & \\ \text{If } \left(\underline{x} \text{ is } HR_j\right) & \text{Then } Out(\underline{x}) \text{ is } w_j \\ \ldots & \\ \text{If } \left(\underline{x} \text{ is } HR_J\right) & \text{Then } Out(\underline{x}) \text{ is } w_J \end{array} \quad (10)$$

where $HR_j \in [m_{j1}, M_{j1}] \times \cdots \times [m_{jp}, M_{jp}]$; the output values are obtained based on the computation of a center average defuzzifier. FHRCNN requires repeated adjustment for each parameter set to achieve optimal classification. First, let each parameter set in FHCRNN correspond to a weight factor located in $[l_1,h_1] \times \cdots \times [l_n,h_n]$, where each vector represents a potential answer to the optimal parameter set. After randomly initializing the vectors and selecting the values for the initial weight vector, $\underline{w_j}(0)$, the winner neuron j^* can be found at time k based on the minimum distance Euclidean criterion; that is:

$$j^* = \underset{1 \le j \le M \times N}{Arg\max} f(w_{j1} \times l_1, \ldots, w_{jn} \times l_n) = \underset{1 \le j \le M \times N}{Arg\max} f(w_{j1} \times 1, \ldots, w_{jn} \times 1)$$

$$= \underset{1 \le j \le M \times N}{Arg\max} f(w_{j1},\ldots,w_{jn}) = \underset{1 \le j \le M \times N}{Arg\max} f(w_j) \tag{11}$$

where $M \times N$ represents the network size, and *f (w)* is the objective function of the optimization problem. In acquiring weight adjustment for j^* and its neighbors, the SOMO algorithm applies:

$$\begin{aligned} &\underline{w}_j(k+1) = \underline{w}_j(k) + \lambda_1 \Lambda_{j^\bullet,j}[\underline{w}_{j^\bullet}(k) - \underline{w}_j(k)] + \lambda_2(1-\Lambda_{j^\bullet,j})\underline{n} \\ &for 1 < j \le M \times N \end{aligned} \tag{12}$$

where

$$\Lambda_{j^\bullet,j} = 1 - \frac{d_{j^\bullet,j}}{d_{\max}} = 1 - \frac{d_{j^\bullet,j}}{\max\left\{d_{j^\bullet,1'}, d_{j^\bullet,N'}, d_{j^\bullet,M\times(N-1)+1'}, d_{j^\bullet,M\times N}\right\}} \tag{13}$$

Both the Λ_1 and Λ_2 are learning rates and the ranges for them are $0 < \Lambda_1 \le 0.3$ and $0 < \Lambda_2 \le 0.2$. $\underline{n} = (n_1, \ldots, n_n)^T$ is the noise vector of the new weight vector. The final step is to perform a certain number or pre-determined number of iterations in an attempt to yield the winner *j** containing the optimal parameter set for FHRCNN.

With the proposed machine learning model based on SOMO and FHRCNN now described, the next section discusses the image processing of bridge cracks.

5. Imaging Processing

The image-processing steps performed in this study are as follows. First, apply image grayscaling to process the original color image. Second, use the high-pass filter to remove the low-frequency noise in the images to highlight the characteristics of a crack. Third, separate the subject and background through the binary process and eliminate unnecessary noise with labeling. Finally, employ MATLAB to develop a Local Directional Pattern (LDP) algorithm to capture crack features and mark the crack part in the original image with red to verify the recognition accuracy.

5.1 Image Identification Preprocessing

The sampling system (e.g. taking lens types, shooting angle) and the spot environment (e.g. light source, uneven illumination) will greatly influence the data obtained from the original image, and thus can contribute to generating image noise. If the image process is conducted without denoising and simplifying, the risk of recognition difficulty or error is increased. Hence, before the binarization of the input image, preprocessing of the original image is necessary. The preprocessing techniques applied in this study include multi-image averaging, spatial domain filtering, frequency domain filtering, and other methods [17].

In this study, the natural light shelter and camera flash were applied before sampling to eliminate the uneven illumination problem, so the pre-processing of the original image is relatively simple. After collecting a large number of images, pre-processing on the computer can be performed to strengthen the information conveyed by the images and to convert them into a more suitable format and type for electronic machine recognition. The exact steps are described as below.

5.2 Grayscale Process

Objects in an image can be easily identified by the human eye, but for electronic machines, the simpler the color and composition of the image are, the quicker the identification of objects will be. The RGB (Red, Green and Blue) three-color parameters compose a color image, with 256 values (0-255) for each parameter to represent different shades. Performing recognition on the original color image will bring redundant computational efforts and reduce operation efficiency.

The purpose of image grayscaling is to convert the color image of 24-bit RGB tricolor ranging from 0 to 255 into a black-and-white image of 8-shade gray value [18]. Hence, the time resource and memory usage space of the machine in operation can be dramatically reduced, which also avoids the redundancy of expression.

Additionally, the original pixel size of the images obtained in this study is 3920x2204, which is too large to be applied during the recognition function and is unfavorable to the identification procedures. Thus the pixel value must be adjusted to approximately one million to be more efficiently assessed by the recognition software. The image grayscale formula is:

$$Y = 0.333\,R + 0.333\,G + 0.333\,B \tag{14}$$

Where, Y is brightness, R is red value, G is green value and B is blue value.

Although the digital images in this research are of concrete surfaces with simple colors, brown dust and green liverworts attached to the structure surface are still represented in the form of 24-bit RGB in color images. Therefore grayscaling remains necessary in this study to store the images in eight shades.

5.3 Filter Processing

A common digital signal processing tool with two significant functions can be called a filter. One function of filtering is to select a specific signal frequency to determine what content will or won't make it through to output. Another function is to suppress the interference of noise. The filtering in this study was primarily applied to suppress the noise in the images, which helps to reduce the sharpness of the grayscale image, smooth edges, eliminate noise, and highlight the crack features

[19].

5.4 Binarization

Binarization is an image segmentation technique. Easy storage, processing, and identification are three advantages of the binary image, which vastly contribute to capturing the specific information of the image in the image recognition processing. Through binary processing, the decreased signal complexity and obvious black-and-white color difference of the image will reduce the error rate in subsequent location, and accelerate the speed of image processing [20].

Binarization generates a dichromatic image, i.e., binary processing partitions the grayscale of an image into two values, which is also called the grayscale threshold. After setting a grayscale threshold (N) in the image, all pixel points in the image are examined. If the grayscale value of the pixel is less than the threshold, let it be a dark point (N=0), otherwise, it is a bright point (N=255). A binary image b(i,j) will be obtained after setting all pixel points, as shown in Eq. 15.

$$b(i, j) = \begin{cases} 1 \, if \;\; f(x, y) \leq N \\ 0 \quad\; otherwise \end{cases} \tag{15}$$

In the binary process, properly setting the threshold is extremely critical. Improper setting will affect the image results. Two methods are often used to obtain the threshold: manual setting and deriving one from the average image grayscale values. Manual setting may cause distortion of the key part, and need to be reset when processing different image. And the result average threshold method in this study is also limited.

This study applies its own alternative method to choose the threshold. Namely, a statistical analysis is performed upon the pixel values to find the grayscale peak value (N), which is set as the value for the binary threshold. Statistics for the pixel values are shown in Figure 1. Repeated tests reveal that if the threshold value of binarization is set in strict accordance with the N values as the peak value in Figure 1, the dark points will be insufficient to clearly display the fracture of the main crack, as too many values just below the threshold are not activated, creating a noncontiguous array of points largely indistinguishable from noise, rather than a distinct crack line. Thus, it was necessary to gradually reduce N values manually in 1 pixel increments. An effective decreased value was identified at 10 pixels, and thus the threshold value was set to be N-10. A binary image is acquired after binarization. However, in addition to highlighting the crack in the concrete surface, noise is expressed in the image as well. Thus, the LABEL is employed to remove noise

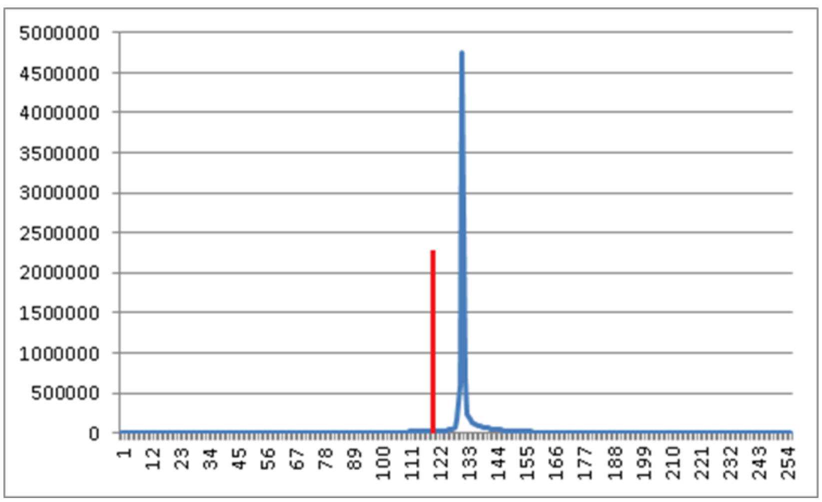

Figure 1. Statistical histogram of grayscale pixel values

5.5 Crack Features Capture

As a result of binary processing, the pixel points of both cracks and non-cracks all have the potential to be assessed as a dark point whose grayscale threshold is 0. To reduce recognition errors, the directivity of the meandering and uneven main crack needs to be used to capture its features. Jabid et al. (2010) put forward the LDP (Local Directional Pattern) algorithm to calculate the gradient values of image edges in different directions and find out the characteristic values of the pixels in various directions. This algorithm is often applied to identify specific edges in an image, which be used for example to automatically distinguish the walking postures between men and women [21], to perform detailed facial expression recognition [22] and perform local face recognition [23]. Even in circumstances with noise or non-monolithic light sources, the gradient values obtained from LDP remain invariant.

The 8-bit binary code of LDP can calculate gradient values of image edges and is encoded to describe the insensitive curve, sideline and corner. The algorithm principle can be interpreted as comparing the gradient values in different directions. In order to obtain gradient values, a Kirsch mask is used to calculate the gradient values of eight surrounding directions (M_0-M7) after randomly selecting a pixel point as a center. Then, the eight corresponding gradient values m0-m7 represent the importance in the relative directions. The image often has a strong reaction in certain directions when analyzed according to this method. After obtaining the eight gradient values, the k main directions need to be determined in order to generate the LDP code. Let the kth large value $|M_j|$ be 1, and the other (8-K) values be 0. The M_k is the kth large value after Kirsch masking. The calculation is shown in Figure 2.

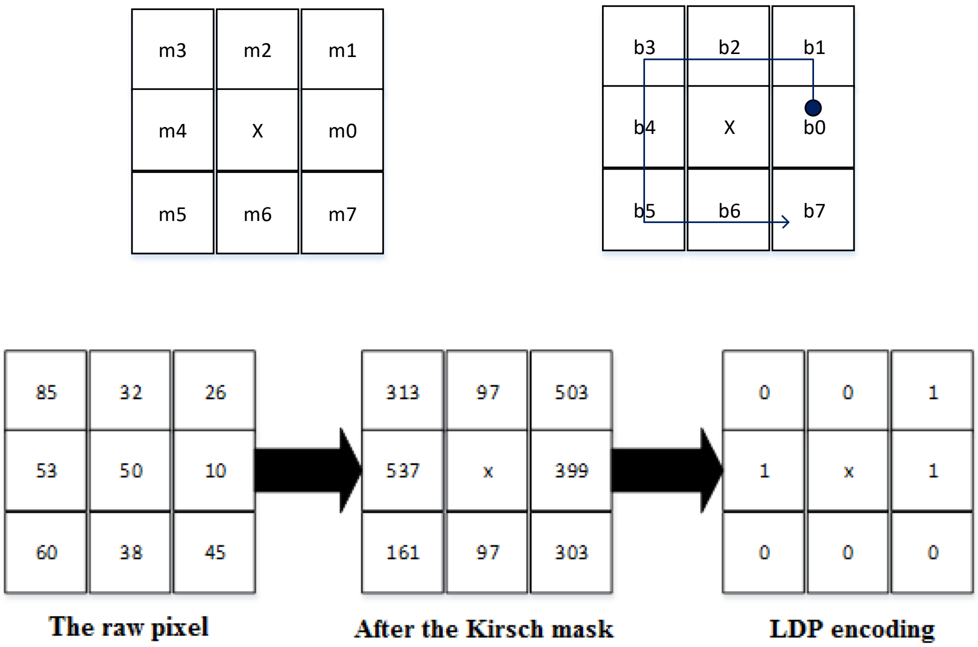

Figure 2. Diagram of LDP algorithm

Since binary processing is already applied to the image sample at an earlier stage, the LDP calculation just needs to aim at the dark point whose threshold is 0 in the figure. In this study, the mask calculation was conducted with respect to the eight directions (east, northeast, north, southwest, northwest, west, south and southeast) stretching out from the center to encode the eight gradient values obtained. The coding method consists of sorting the eight directions according to their absolute value, then setting the value of the first three directions as 1 and the last 5 as 0. As a result, the main crack line with bending directionality can be steadily strengthened, while the noise without directions is removed.

5.6 Direction Detection of Objective figure

Images processed according to the above procedures are then ready to be recognized by machine vision. This paper employs the imaging direction detection method proposed by Stock and Swonger to obtain the block directional image of the cracks, i.e. segmenting the crack into several square blocks. In each block, the main direction in the crack block is calculated using local crack line information (i.e. the grayscale values obtained from the image processing steps). The calculation method needs to employ a 9 x 9 mask, as shown in Figure 3. Let S_i to be the sum of pixel label i's grayscale values in the mask, where $i = 0, 1, \ldots, 7$. S_p refers to the minimum of all the values of S_i, while S_q means the maximum among the values of S_i, where, $p, q = 0, 1, \ldots, 7$. C represents the grayscale value of the mask center. The d is the final determined direction. Then:

$$d = \begin{cases} p \; if \; (4C + S_p + S_q) < \frac{3}{8} \sum_{i=0}^{7} S_i \\ q \qquad\qquad otherwise \end{cases} \tag{16}$$

6		5		4		3		2
7		6	5	4	3	2		1
		7				1		
0		0				0		0
		1				7		
1		2	3	4	5	6		7
2		3	17	4	3	5	23	6

Figure 3. The 9×9 mask applied by Stock and Swonger [24]

An advantage of the method is that eight directions can be detected, while a weakness is that the size of the mask is fixed, which may lead miscalculation from a small amount of noise determined by four grayscale values in each direction. For the purpose of increasing the accuracy rate of the recognition technique, 36 images were randomly selected from the set of 216 for training and the other 4 comprising the recognition set. The red line marked along the crack of original image is regarded as the training target and used to control accuracy. Then the red-marked cracks were set with the desired outputs (crack is 1, non-crack is 0), and FHRCNN was employed to perform the classification. After conducting LDP calculations on each image, the 9×9 mask was used. The input data is the representative surrounding points whose LDP label is more than 0. Thus, the dimension size of the input data is 81 (as shown in Eq. 16).

This study randomly selected 18 images to be the training data set, and each image has approximately 1 million points. 5000 stochastic points are chosen to be representative points. Hence the FHRCNN recognition rates are obtained, where the training accuracy rate is 100% and the test accuracy rate is 81%. Two values are generated in the results: true positive and false positive, which refer to the recognition accuracy rate of crack points and non-crack points respectively. Both of the values will affect the judgment of the crack's state. During the training process, strong light, less sharp cracks, excessively thin crack lines, and noise will all influence the accuracy rate. Therefore, considering that not all images meet the research requirements, clean images without too much noise are chosen for training.

6. Case Implementation

In order to verify that the recognition program can be used for practical bridge inspection, the case study executed in the present research is the severely damaged Hsichou Bridge (shown in Figure

4.) selected from the Taiwan Bridge Management System. Manual shooting was conducted in the field and then qualified images were selected to perform the recognition test. An example of an image analyzed for this case study is shown in Figure 5.

Figure 4. Hsichou Bridge

Figure 5. Image of cracks on the Hsichou Bridge

50 pictures of the cracks on the Hsichou Bridge were taken for this study. 36 pictures were selected according the previously described filtering rules used to verify the accuracy rate. The process is presented below. After the grayscale and high-pass filtering processes, the image shown as figure 6 is converted as shown in figure 7.

The peak value corresponding to the grayscale value of the Figure 7 was calculated with the pixel statistic tool. A grayscale peak value of 133 was obtained, as shown in Figure 8. Then 122 (133-10=122) was set as the binarization threshold to conduct binary processing and obtain the image shown as Figure 9. However, except for the crack, numerous black miscellaneous points need to be removed using Labeling. This step is followed by applying the LDP algorithm to highlight the directional characteristics of the crack. In this way, the image shown as Figure 10 can be obtained.

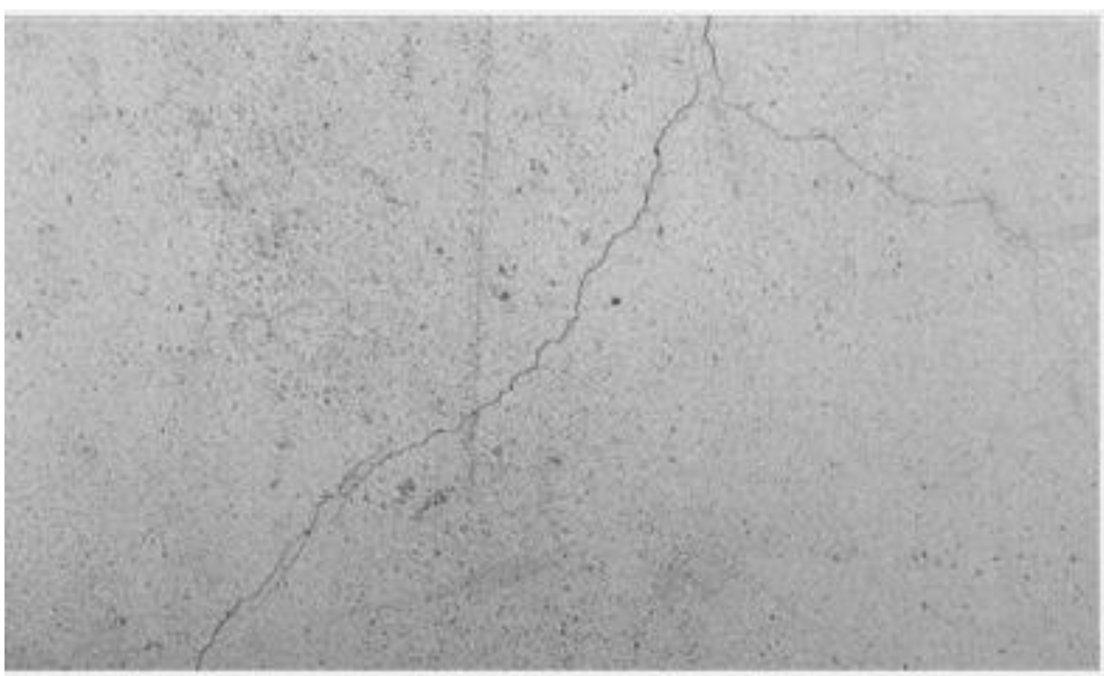

Figure 6. The image of the cracks of Hsichou Bridge

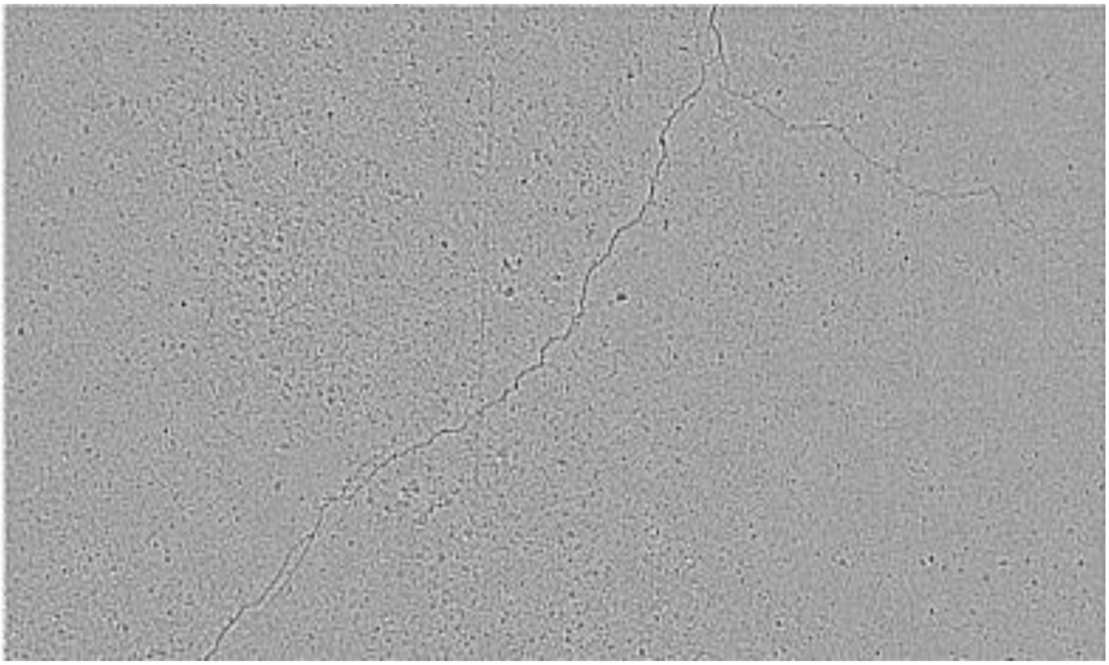

Figure 7. The image after the grayscale and high-pass filtering processes

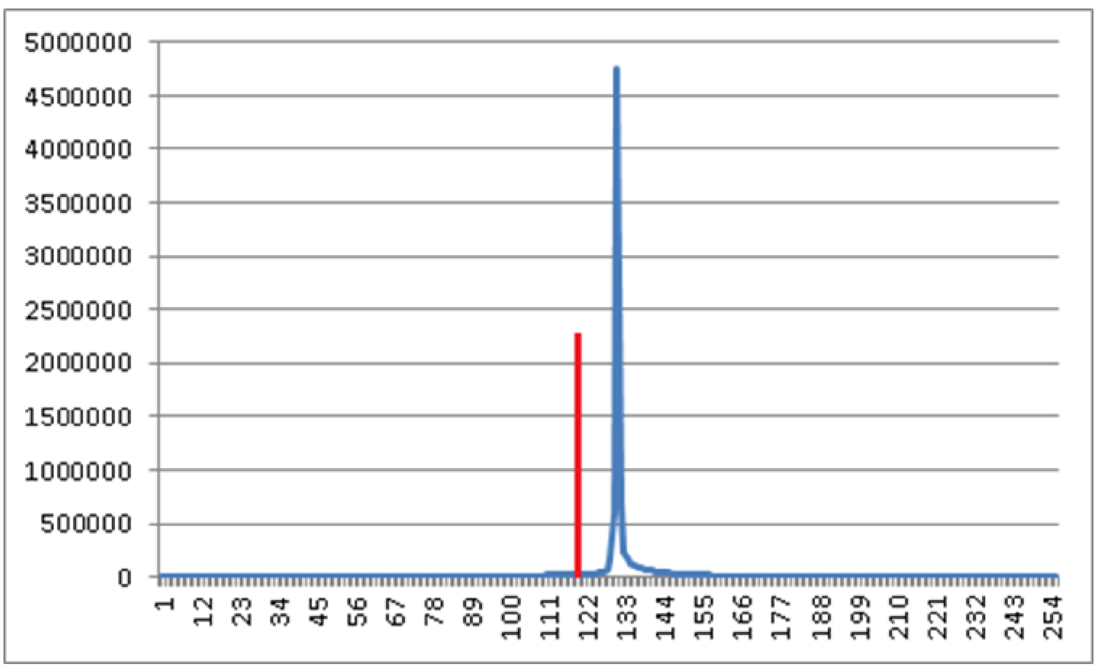

Figure 8. Grayscale histogram for Figure 5.9

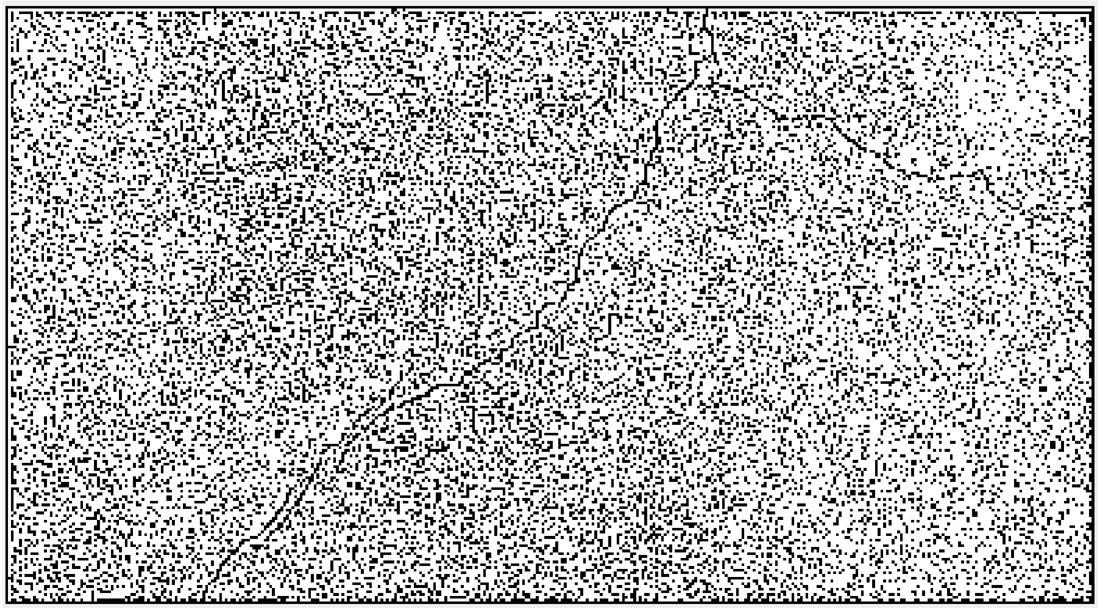

Figure 9. The image after the binarization

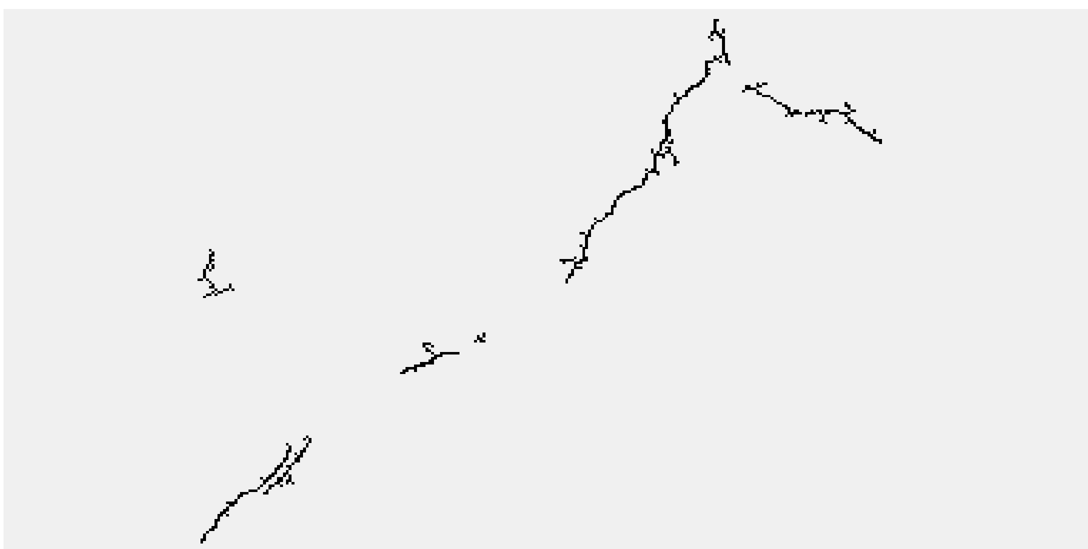

Figure 10. The image after the LDP algorithm

The next step is to mark the red line along with the crack in the original image as shown in Figure 11. Note that the red line must cover the main crack line to ensure the recognition results.

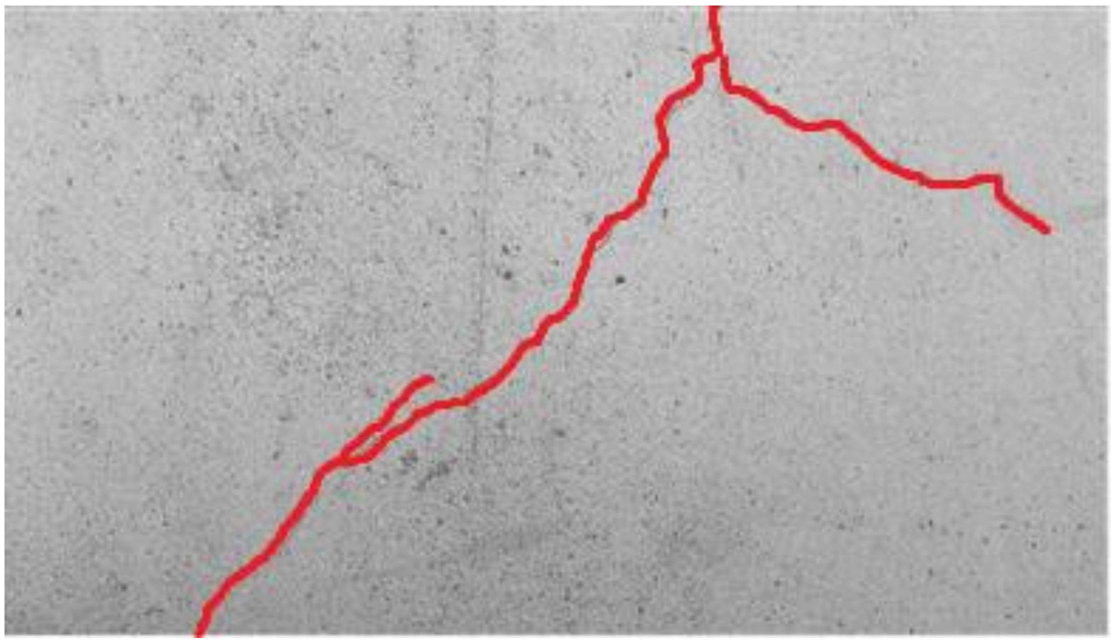

Figure 11. The original image with red line

Figure 12. The crack recognition result

The recognition results: True positive: 0.893; False positive: 0.909

In the image shown in Figure 12, the recognition rate of the points representing the crack is 89%, while that of the points referring to the non-crack is approximately 91%. After performing the statistical treatment on all samples of the Hsichou Bridge, the success rate of the recognition program researched and developed by this study is 89% or more, which is 8% higher than the accuracy of the test sample group (81%). Thus this recognition program can effectively and reliably identify concrete bridge cracks. Other bridge crack detection technologies have image recognition accuracy rates averaging between 74% to 96%, depending on the different image types, training set types, and the feature extraction methods [4-8] used. Thus the results of the current study demonstrate performance comparable to other state-the-art methods tested in previous studies.

7. Conclusions

Though visual inspection of bridges in Taiwan is relatively economical, this laborious and time-consuming method is easily influenced by the subjective behavior of individual inspectors. The automatic image recognition technique developed by this paper is aimed at effectively and efficiently identifying the cracks in concrete bridges using machine vision, rather than traditional human judgment, substituting subjective efforts with objective testing results.

This study selected appropriate bridges from the Taiwan Bridge Management System (BMS) and selected samples of appropriate images taken by a hand-held digital camera. To control the intensity of the illumination parameters, this study used artificial light as the only light source in the shooting process by shielding the cracks from natural light. Further image processing was performed in this study to ensure that the ultimately transformed color images could be quickly analyzed by the machine for crack recognition. which is described below. To do so, the color image is first converted to the 8-bit grayscale image through grayscaling. Then a high-pass filter is applied to strengthen the crack edge in the image. Next the grayscale image is converted into simpler black-and-white image through binarization. The fourth step removes the noise with labeling. Fifthly, crack characteristics are extracted through the Local Directional Pattern (LDP) algorithm. Finally, training to obtain a satisfactory recognition rate was performed using the FHRCNN classification method, which would result in one of two values, true positive and false positive, which represent the accuracy rates in recognizing crack and non-crack points respectively. Quantifying the crack dimensions can be also achieved if the photo scale is set to be a constant shooting distance.

The accuracy rate reached up to 81% after training the recognition program in this research. In the case study, the accuracy rate in recognizing the Hsichou Bridge cracks reached 89%. In another words, the results of this research demonstrate that the computer can successfully recognize the cracks through converting the image data into numerical values. The high identification ability of the recognition program in this study ensures it would have great credibility when applied in practice. The standardized computation of the recognition results is also able to overcome the inconsistency in applying criteria that arises from subjectivity of human visual judgment. When severe natural disasters would otherwise especially necessitate bridge inspections by a great deal of proficient inspectors, compared with other traditional artificial inspection methods, the image recognition program combined with the image acquisition criteria featured in this study is more economical and efficient. Since the initial onsite inspection would only require the use of a digital camera and light shelter, the time and cost in training inspectors to use such equipment can be immensely reduced.

Nevertheless, the recognition technique in the original images only directs to adjust the field conditions to highlight the cracks. The surfaces of many bridges, especially those affected by natural disasters or construction, the dirt, stains, and construction character markings can contribute to the misrecognition of cracks. As for the artificial light, though the unnecessary reflected light is absorbed by the black paper in this study, the hard light can wipe out some tiny cracks, resulting in increased recognition difficulty. Thus future research could apply artificial lighting that is more adjustable in intensity to avoid crack disappearance. A follow-up study could focus on improving the image processing to overcome the extra effects beyond the noise. Though the scope of this paper’s research is limited to concrete bridges, after adjusting the relative parameters in the same image processing, this recognition technique could also be applied to other concrete buildings (e.g. general houses, walls, roads, etc.). In addition, this technique currently can only judge whether the crack exists or not. If the data of various crack shapes and characteristics can be collected to set up a crack database, the recognition efficiency would be greatly enhanced. Furthermore, as the initial acquisition of the images for this paper mainly relies on manual work, the steep ruggedness of the terrain and fast-flowing rivers at the inspection sites all bear upon the safety of the inspectors. If a fully-automated camera system (e.g. a remote machine for shooting the images) can be developed to work with the image recognition technique presented in this paper, some bridges that are difficult for inspectors to reach can be successfully inspected. To obtain crack directionality, this study employs the Stock and Swonger calculation method. A limitation of this method, however, is that a small amount of noise may generate misjudgment. In addition, the influence of the single reference point is quite large, considerably increasing the possibility of misrecognition. Thus, future research could explore finding a method to reduce calculation error based on the Stock and Swonger calculation method, such as using smoothing processing to smooth the curve and angle. Future research could also further investigate the effects of different image gray-scale distributions on the prediction results.

References

[1] Federal Highway Administration (FHWA), Specification for the national bridge inventory bridge elements, January, 2014.

[2] Federal Highway Administration (FHWA), Long-term bridge performance primer, Rep. No. FHWA-HRT-13-051, U.S. Dept. of Transportation, Washington, D.C. December, 2013.

[3] Li, G., He, S., Ju, Y., and Du, K. Long-distance precision inspection method for bridge cracks with image processing, Automation in Construction, 41, 83-95, 2014. DOI:10.1016/j.autcon.2013.10.021

[4] Yeum, C.M. and Dyke, S.J. (2015). Vision-Based Automated Crack Detection for Bridge Inspection, Computer-Aided Civil and Infrastructure Engineering 30, 759-770. DOI: 10.1111/mice.12141

[5] Oh, J. K., Jang, G., Oh, S. Bridge inspection robot system with machine vision, Automation in Construction, 18(7): 929-941, 2009. doi:10.1016/j.autcon.2009.04.003

[6] Zhu Z, German S, Brilakis I. Detection of large-scale concrete columns for automated bridge inspection, Automation in Construction, 19(8), 1047-1055, 2010. DOI:10.1016/j.autcon.2010.07.016

[7] Yu., Q., Guo, J., Wang, S., Zhu, Q., and Tao, B. (2012). Study on a New Bridge Crack Detection Robot based on Machine Vision, Intelligent Robotics and Applications, Volume 7506 of the series Lecture Notes in Computer Science, 174-184.

[8] Bu, G.P., Chanda, S., Guan, H., Blumenstein, M., and Loo, Y.C. (2015). Crack Detection using a Texture Analysis-based Technique for Visual Bridge Inspection, Special Issue: Electronic Journal of Structural Engineering 14(1), 41-48.

[9] Li, Y. F., Hsieh, S. H., and Chan, C. C. A planning multi-level bridge management system for Taiwan, Automation and Robotics in Construction, XVII: 71-75, 2000.

[10] Adhikari, R. S., Moselhi, O., Bagchi, A. Image-based retrieval of concrete crack properties for bridge inspection, Automation in Construction, 39: 180-194, 2014. DOI:10.1016/j.autcon.2013.06.011

[11] Konietzky, H. ed. Numerical modelling of discrete materials in geotechnical engineering, civil engineering and earth sciences: Proceedings of the First International UDEC/3DEC Symposium, Bochum, Germany, 29 September-1 October 2004. CRC Press, 2004.

[12] Taiwan Highway Administration, The manual of highway maintenance, 2012.

[13] Su, M. C., Zhao, Y. X., and Lai. E. SOM-based Optimization, In Proceedings of IEEE international joint conference on neural networks, Budapest, Hungary, 781-786.

[14] Chen J. H., Su-M. C., Zhao, Y. X., Hsieh, Y. J., and Chen, W. H. Application of SOMO based clustering in building renovation. International Journal of Fuzzy Systems, 10(3), 195-201, 2008.

[15] Chen, J. H., Su, M. C., and Yang, L. R. Comparison of SOM-based optimization and particle swarm optimization for minimizing the construction time of a secant pile wall. Automation in Construction, 18(6), 844–848, 2009. DOI:10.1016/j.autcon.2009.03.008

[16] Chen, J. H. A hybrid knowledge-sharing model for corporate foreign investment in China’s construction market, Expert Systems with Applications, 39, 7585-7590, 2012. DOI:10.1016/j.eswa.2011.11.076

[17] Garris, M., Wilson, C., and Blue, J. Neural network-based systems for handprint OCR applications, IEEE Transactions of Image Processing, 7(8), 1097-1112, 1998. DOI: 10.1109/83.704304

[18] Bettahar, S., Stambouli, A. B., Lambert, P., and Benoit, A. PDE-based enhancement of color images in RGB space, IEEE Transactions on Image Processing, 21(5), 2500 – 2512, 2012. DOI: 10.1109/TIP.2011.2177844

[19] Bazi, Y. and Melgani, F. Gaussian process approach to remote sensing image classification, IEEE Transactions on Geoscience and Remote Sensing 48, 186-197, 2010. DOI: 10.1109/TGRS.2009.2023983

[20] Jin L. and Li D. A switching vector median filter based on the CIELAB color space for color image restoration, Signal Processing, 87(6), 1345-1354 , 2010. DOI:10.1016/j.sigpro.2006.11.008

[21] Jabid, T., Kabir, M. H., and Chae, O. Gender classification using local directional pattern (LDP), International Conference on Pattern Recognition, 2162-2165, 2010.

[22] Kabir, H., Jabid, T., and Chae, O., Local directional pattern variance (LDPV): A robust feature descriptor for facial expression recognition, International Arab Journal of Information Technology, 9(4), 382- 391, 2012.

[23] Castillo, J. A. R., Rivera, A. R., and Chae, O., Robust facial recognition based on local Gaussian structural pattern, International Journal of Innovative Computing, Information and Control, 8(12), 8399-8413, 2012.

[24] R. M. Stock, C. W. Swonger, Development and Evaluation of a reader of Fingerprint Minutiae, Cornell Aeronautical Laboratory, Technical Report CAL XM-2478 2478 -X-1:13-17, 1969.