DAMAGE ASSESSMENT OF RC SLABS USING A PRACTICAL ENSEMBLE SYSTEM FOR MULTIPLE FEATURE SELECTION

Koichiro Nakatsu
Kansai University, Japan

Hitoshi Furuta
Kansai University, Japan

Yasutoshi Nomura
Ritsumeikan University, Japan

Kyosuke Takahashi
Kansai University Graduate School, Japan

Ken Ishibashi
Kansai University Graduate School, Japan

Yuji Hayakawa
OBIC CO., Ltd., Japan

Recently, it has become more and more important to establish a rational maintenance program for infrastructures. In bridge structures, Reinforced Concrete (RC) slabs are suffering from damage, because they directly resist applied loads. The damage of RC slabs is assessed with the aid of digital photos and pattern recognition. In this study, an attempt is made to develop a new damage detection method for RC slabs with cracks, by using pattern recognition techniques. In the pattern recognition, the feature extraction is a quite important process, because the recognition rate depends on the characteristic of the recognition object. However, it is difficult to select the appropriate feature for the recognition systematically. The previous research proposed the pattern recognition method using the ensemble system based on fuzzy classifier for multiple feature selection. The previous method uses classifiers based on each feature and each pattern recognition algorithm. However, this method is not suitable for the problems with many input vectors because it takes a lot of time for learning when the number of the input vector increases. The proposed method can overcome the problem by introducing ID3 (Interactive Dichotomizer 3) with classifiers consisting of many feature vectors. In the proposed method, ID3 constructs the decision tree for the damage assessment with the results obtained from classifiers based on each feature and each pattern recognition algorithm. Using the decision tree, it is possible to select appropriate classifiers among many prepared classifiers as less as possible so as to find an ensemble system with higher performance and less computation load. Numerical examples are presented to demonstrate the efficiency and applicability of the proposed method.

Corresponding author's email: inside2@sc.kutc.kansai-u.ac.jp

DAMAGE ASSESSMENT OF RC SLABS USING A PRACTICAL ENSEMBLE SYSTEM FOR MULTIPLE FEATURE SELECTION

Koichiro Nakatsu[1], Hitoshi Furuta[1], Yasutoshi Nomura[2], Kyosuke Takahashi[3], Ken Ishibashi[3], and Yuji Hayakawa[4]

[1] Faculty of Informatics, Kansai University, Takatsuki, Japan
[2] Faculty of Engineering, Ritsumeikan University, Kusatsu, Japan
[3] Graduate School of Informatics, Kansai University, Takatsuki, Japan
[4] OBIC Co., Ltd., Chuo, Japan

ABSTRACT: Recently, it has become more and more important to establish a rational maintenance program for infrastructures. In bridge structures, Reinforced Concrete (RC) slabs are suffering from damage, because they directly resist applied loads. In this study, an attempt is made to develop a new damage detection method for RC slabs with cracks, by using pattern recognition techniques. In the pattern recognition, the feature extraction is a quite important process, because the recognition rate depends on the characteristic of the recognition object. The previous research proposed the pattern recognition method using the ensemble system based on fuzzy classifier for multiple feature selection. However, this method is not suitable for the problems with many input vectors because it takes a lot of time for learning when the number of the input vector increases. The proposed method can overcome the problem by introducing ID3 (Interactive Dichotomizer 3) with classifiers consisting of many feature vectors. Numerical examples are presented to demonstrate the efficiency and applicability of the proposed method.

1 INTRODUCTION

Recently, it has become more and more important to establish a rational maintenance program for infrastructures. In bridge structures, Reinforced Concrete (RC) slabs are suffering from damage, because they directly resist applied loads. In order to establish a rational management program for bridge structures, it is necessary to evaluate the structural damage of existing bridges in a quantitative manner. The damage of RC slabs is, in general, assessed through the visual inspections. However, it is difficult to avoid the subjectivity of inspectors when visual data are used for the evaluation of damage or deterioration. In this study, an attempt is made to develop a new damage detection method for RC slabs with cracks, by using pattern recognition techniques. In the pattern recognition, the feature extraction is a quite important process, because the recognition rate depends on the characteristic of the recognition object. However, it is difficult to select the appropriate feature for the recognition systematically. The previous research proposed the pattern recognition method using the ensemble system based on fuzzy

classifier for multiple feature selection (Furuta et al. 2009). The previous method uses classifiers based on each feature and each pattern recognition algorithm. However, this method is not suitable for the problems with many input vectors because it takes a lot of time for learning when the number of the input vector increases. The new method proposed in this paper can overcome the problem by introducing ID3 (Interactive Dichotomizer 3) (Quinlan 1086) with classifiers consisting of many feature vectors. In the proposed method, ID3 constructs the decision tree for the damage assessment with the results obtained from classifiers based on each feature and each pattern recognition algorithm. Using the decision tree, it is possible to select appropriate classifiers among many prepared classifiers as less as possible so as to find an ensemble system with higher performance and less computation load. Numerical examples are presented to demonstrate the efficiency and applicability of the proposed method by using the actual data.

2 PATTERN RECOGNITION FOR DAMEGE ASSESSMENT

Damage assessment by inspectors is generally carried out from visual data paying attention to about the factors that show apparent damage such as cracks in concrete surface. Then, it is difficult to avoid the subjectivity of inspectors when visual data are used for the evaluation of damage or deterioration. Therefore, it is useful to develop a new quantitative damage detection method for concrete slabs with cracks. In order to achieve the objective quantitative evaluation, pattern recognition technologies have been utilized to evaluate the damage of RC slabs with cracks by combining the data given by digital photos.

2.1 Application of ensemble system

In the pattern recognition, the feature extraction is quite important to control the recognition performance. However, it is difficult to achieve the generic feature extraction method because of the dependence on the recognition object. In the damage assessment of RC bridge decks, the feature extraction is also extremely important. When the feature extraction from digital photos of RC bridge decks is conducted, there are various features such as continuity or concentration of the width of the crack. It is difficult to use all features for the recognition because of the problem called “curse of dimensionality” (Bellman 1961). Thus the specific feature is selected among various features and it is used for the recognition. However, it is very difficult to select the most effective feature among various features. Even if the most effective feature is extracted, there is the limit in recognition rate using the feature. Therefore, it is necessary to construct the system that selects the feature effectively and automatically. One of the problems of pattern recognition is the relationship between the problem to be solved and the performance of various pattern recognition algorithms. However, the recognition performance is different according to the combination of the algorithm and the features that are extracted from the recognition object. Therefore, it is necessary to select an appropriate algorithm considering the recognition object and the feature. To solve these problems, we proposed a damage assessment method by introducing fuzzy ensemble learning method (Nakashima et al. 2003) with classifiers consisting of many feature vectors (Furuta et al. 2009). The classifier ensemble system in the previous method consists of fuzzy classifiers or neural network classifiers and fuzzy ensemble. Fuzzy classifiers and neural network classifiers are constructed according to each feature. The fuzzy ensemble estimates weight value to each classifier result. The best classifier result

is selected from classifiers result using this weight value. The results of applying the previous methods, we were able to demonstrate the practical system for damage assessment system.

2.2 Problems of pattern recognition for damage assessment

The previous method uses classifiers based on each feature and each pattern recognition algorithm (Furuta et al. 2009). However, this method is not suitable for the problems with many input vectors because it takes a lot of time for leaning when the number of the input vector increases. The fuzzy ensemble system uses a fuzzy classifier to select classifier systems from an input pattern. Therefore, increase in the number of dimensions of the input vectors leads to an exponential increase in the number of rules used in fuzzy classifier. Recent study has succeeded to minimize the number of if-then rules without loss of accuracy using optimization techniques, such as Genetic Algorithm (GA) (Ishibuchi and Yamamoto 2004). However, there is a difficulty to reduce the number of rules for the case with unknown values. In order to apply such an optimization method as GA, there must be an evaluation function to minimize the number of if-then rules considering the generalization performance. Moreover, even if it could reduce the number of if-then rules, it would cause problems in terms of usability and the determination of design parameters for minimizing the number of if-then rules. In this study, we attempt to build the damage assessment system which can select appropriate classifiers for the case with large input dimensions without reducing the number of fuzzy if-then rules.

3 PROPOSED METHOD

The proposed method can overcome these problems by introducing ID3. The outline of the proposed method will be explained in the following.

3.1 Interactive Dichotomizer 3

ID3 is a supervised learning algorithm to classify the data based on decision trees. A decision tree is shown in Figure 1, which is generated through data sets presented in Table 1 by ID3.

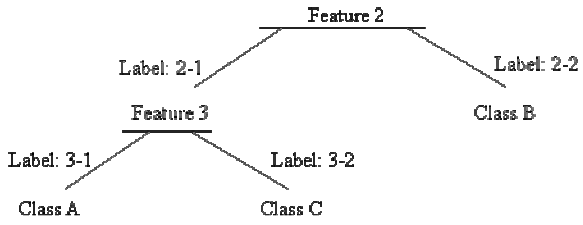

Figure 1. A decision tree ID3.

Table 1. Data sets.

Feature 1	Feature 2	Feature 3	Class
1-1	2-1	3-1	A
1-1	2-2	3-1	B
1-2	2-2	3-1	B

1-1	2-1	3-2	C
1-2	2-1	3-1	A

The following shows the flow of classification of data in Table 1. First, the value of "Feature 2" located at the top node is checked. And by tracing the links according to their values, the attributes to be next checked can be identified. By repeating this process, it is possible to know the class to be categorized when the search reached the leaf node that has no links. ID3 is implemented using the concept of information entropy. Figure 1 is obtained by using the expected value of information entropy gained on the three attributes shown in data set (Table 1). In this example, "Feature 1" is not used for the classification. Thus, the ID3 decision tree is generated by the minimal information. It has the advantage that it can recognize the process of making the decision tree and classify the data using extra attributes.

3.2 ID3 Ensemble System

The structures of ID3 ensemble system is shown in Figure 2. As well as ID3, the proposed ID3 ensemble system constructs a decision tree to classify a set of learning data. In order to achieve the appropriate classifier from multiple classifiers, the proposed ID3 ensemble system is constructed as shown in Figure 2.

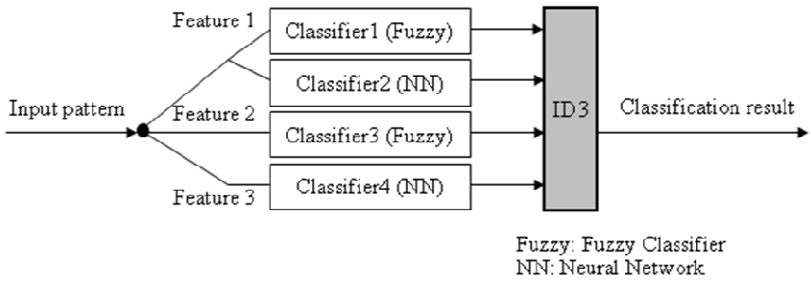

Figure 2. ID3 ensemble system.

The set of learning data for ID3 ensemble system is the recognition result of each classifier. Pattern recognition method used in the proposed method generates a decision tree for these data. This is the decision tree classification based on the evaluation results given by the inspectors. Since the proposed method does not use a vector image pattern from the learning data, the number of attributes used for making the decision tree is the same as the number of classifiers regardless of the feature and classifier selected. Therefore, the ID3 ensemble system enables to deal with multi-dimensional input data. In addition, it is possible to trace the process of data recognition from the decision tree. This means that it is possible to automatically select an appropriate classifier without sacrificing the benefits of the existing system. An example of a decision tree generated from Figure 2 is shown in Figure 3.

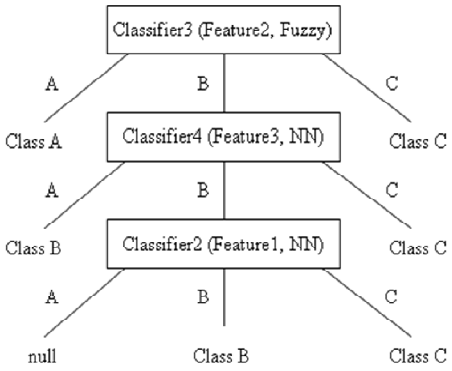

Figure 3. An example of decition tree generated ID3 ensemble system.

In this example, the ID3 ensemble system produces the system with four classifiers, however, there is a case that the set of learning data is not used for recognition classifier shown in Figure 3 (Class 1), because it generates a decision tree to classify all the data with minimal information. Thus, there are both cases that a classifier can classify and multiple classifiers are necessary to classify depending on the characteristics of classes. In Figure 3, “null”, in the decision tree means that the input patterns can not be classified. If the classification of “null” occurs, it means that there is no data among the learning data. If there is a sufficient amount of learning data, the case of “null” is rare. Nevertheless, “null” is also possible to compensate for a pattern to evaluate the data and such data can be involved in the learning data through the detail analysis. This process can contribute to improve the recognition rate.

4 DAMEGE LEVEL DETECTION OF RC BRIDGE DECK

In this study, the damage of RC bridge decks is evaluated with the aid of digital photos and pattern recognition. In general, the procedure for extracting the characteristics of cracks showing up on concrete decks through digital image and the classification of the damage level based on the characteristics are used in the typical pattern recognition system.

4.1 Preprocessing

To obtain the test material, digital images of RC decks taken by a digital camera used. The total number of digital images is 46 and each image is converted to the grayscale image. The digital images used in this study are obtained by marking the cracks with chalk. The damage levels for all digital images are classified into three categories by the inspectors.

Cracks in a digital image of concrete deck are detected in accordance with the following procedure (Figure 4). After all of these processing, the crack pattern is obtained and used for extracting the characteristic of digital images.

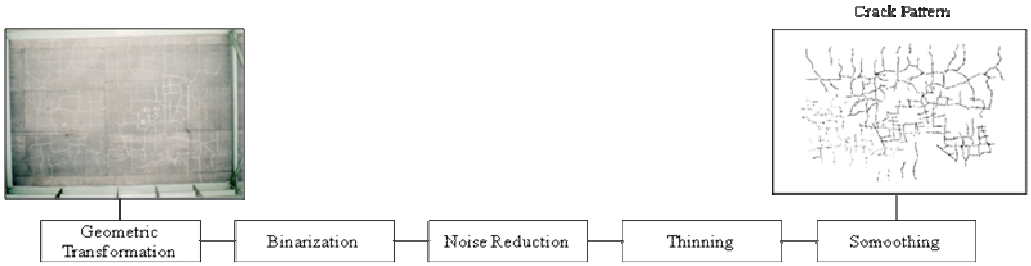

Figure 4. Procedure of image pre-processing.

4.2 Feature extraction

Feature extraction of cracks is conducted from binary images. However, binary images are compressed to 50 × 50 pixels rectangle because original binary images have 1782 × 1220 pixels, take a lot of time to construct classifiers. The features that are extracted from the binary images are the density of cracks, the average and variance of each part and the cross pointing and branch pointing (Figure 5).

In the density of cracks, black pixels are searched for from top left to bottom right of the binary images and the box of 7 × 7 pixels is made centering on the black pixels. The number of black pixels in this box is used as the feature of density of cracks. In the average and variance, the average value and the variance value of the number of black pixels in each area where the binary image was divided into the cross are used. In the cross pointing and branch pointing, the numbers of cross point and branch point are used.

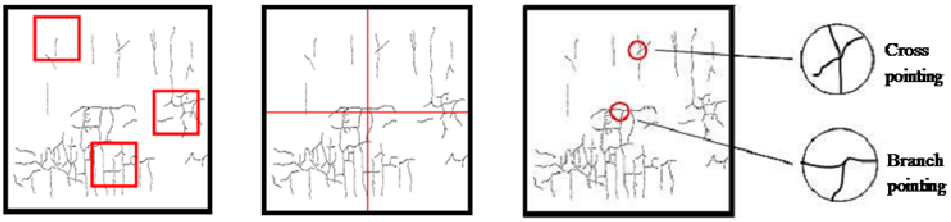

Figure 5. Extracted features of binary image.

4.3 Construction of the system

In this study, the classifiers are constructed by three features and three recognition algorithms. Three features are density of cracks, average and variance and cross pointing, and three recognition algorithms are fuzzy logic, neural network and support vector machine. In the fuzzy classifiers, triangular fuzzy sets are used and attribute is divided into 3 and 4 fuzzy sets. Thus, in each feature, four classifiers are constructed, and twelve classifiers are constructed totally. The ID3 ensemble system generates the decision tree to each classifier result.

Recognition learning of each classifier, 9 images (3 images in each class) are used as learning data among 46 images. The other 37 data are used to test the classifier. ID3 ensemble system, the recognition results as inputs for each classifier. Therefore, in recognition of 9 results for each classifier learning is to be recognized by almost all the results and it has a simple generalization performance for unknown data are generated

by this decision tree. In this study, recognition learning of the ID3 ensemble system, 37 images (different learning data of each classifier) are used as learning data.

4.4 Recognition performance

In the numeral experiments, learning data for the ID3 ensemble system are used while maintaining the ratio for each level of class, 20%-80% by varying the rate selected randomly for learning. Table 2 shows the recognition rate of the test data of each classifier for three features. Table 2 shows the recognition rate of the trial data of each classifier and Table 3 shows the results by the ID3 ensemble system. The results in Table 3 mean the average values obtained by five trials for each trails and the recognition rates of the recognition rates for each damage level and overall.

Neural network classifier using the cross pointing and branch pointing as the feature obtained the highest performance as 84% in the classifiers. Comparing the results of twelve classifiers, it can be confirmed that there are differences at the recognition rate by the difference of the features even with the classifiers using the same algorithm. Although the ID3 ensemble system provides lower recognition rate, this is due to the fact that the learning data are different from the results by each classifier. Namely, the proportion of learning data and unknown data is different so that the proposed method is not inferior to the individual classifier. In addition, Table 3 shows that the result of 80% leaning data greatly varies by the difference because the increase of number of learning data occur the decrease of unknown data. Experimental results presents that while changing the percentage of learning data the proposed method shows the

Table 2. Recognition result.

Feature	F3	F4	NN	SVM
Density (Cracks)	59%	78%	81%	70%
Average-Variance (Cross division)	65%	49%	57%	19%
Cross pointing Branch pointing	70%	81%	84%	68%

F3 : Fuzzy classifier (3 division Membership function)

F4 : Fuzzy classifier (4 division Membership funciton)

NN : Neural network classifier

SVM : Support vector machine with polynomial kernel function classifier

Table 3. Recognition result by ensemble system.

Rate of supervised data (number)	Recognition Rate			
	A	B	C	Total
20% (9/37)	95%	75%	73%	76%
40% (16/37)	93%	80%	62%	72%
60% (23/37)	100%	90%	85%	89%

80% (30/37)	100%	50%	95%	83%

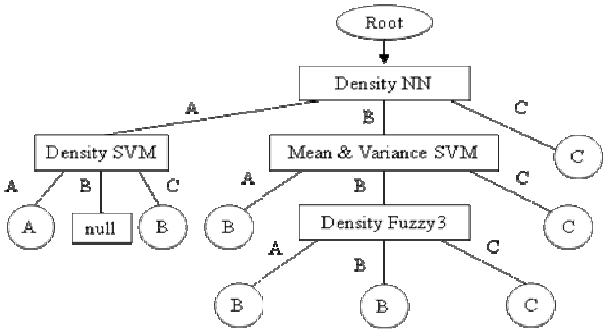

Figure 6. An example of decision tree.

highest recognition rate when the proportion is 60% and the recognition rate decreases when the proportion becomes more than 60%. Then, if provided sufficient data can be used to recognize proper classification can be realized. As details of the proposed method for the damage assessment, Figure 6 shows the decision tree is generated by using 37 images.

Figure 6 shows that even if an image classified as “A” by the neural network classifier using the feature of the density is not output as the decision result, it is needed to re-check it by SVM classifier using the feature of the density. In this case, if SVM using the density provides the results of “A”, the final decision becomes “A”, and if SVM classifies “C”, the final decision is determined as “B”. The reason is due to the fact that the images classified as “B” do not have distinct boundaries to the class “A” or “C”. In fact, “B” appears many times in the decision trees. While the previous method can not classify the damage level “B”, the propose method can provide proper decisions on the classification of damage level.

5 CONCLUSIONS

In this study, an attempt was made to develop a new damage assessment system for RC bridge decks by using the pattern recognition technique and visual information provided by digital photos. The proposed system can automatically select appropriate features that are quite important to classify the damage state into several categories. In order to improve the recognition performance, ID3 was introduced into the damage assessment with such classifiers as ensemble systems, neural networks, and fuzzy reasoning. From the numerical experiments, it is confirmed that the recognition result of these classifiers are quite different. By using the ID3 ensemble system proposed in this paper, it is possible to select the appropriate classifiers effectively from prepared classifiers and to improve the recognition performance efficiently. Though several numerical experiments, the following conclusions are derived:

1. In this study, we proposed ID3 ensemble system for the damage assessment of RC bridge decks. By using the proposed system, it is possible to derive a decision tree to minimize the number of decision rules important for the damage assessment. Since the proposed method does not use a vector image pattern from the learning data, the number of attributes used for making the decision tree is the same as the number of classifiers regardless of the feature and classifier selected. Therefore, the proposed

ID3 ensemble system enables to deal with multi-dimensional input data. In addition, it is possible to trace the process of data recognition from the decision tree.

2. In order to establish a practical damage assessment system, it is necessary to evaluate the damage state based on the damage image by considering many damage factors such as crack width, length and density corresponding to a number of features. Since the proposed system can deal with multiple features, it is useful for practical use as well as the judgments by experienced inspectors.

3. In the future, more improvement of recognition performance can be desirable by introducing more features and algorithms into the ensemble classifier system.

6 REFARENCES

Bellman, R. (1961). "Adaptive Control Processes", *Princeton University Press*.

Furuta, H., Nakatsu, K., Hattori, H., Adachi, W. and Nomura, Y. (2009). "Damage Assessment of Reinforced Concrete Bridge Decks Using Fuzzy Ensemble System", *Proceedings of the 10th International Conference on Structural Safety and Reliability, ICOSSAR*, 3956-3963.

Ishibuchi, H. and Nakashima, T. (2001) "Effect of Rule weights in Fuzzy Rule-Based Classification Systems", *IEEE Trans. on Fuzzy Systems*, 9(4), 506-515.

Ishibuchi, H. and Yamamoto, T. (2004). "Fuzzy Rule Selection by Multi-Objective Genetic Local Search Algorithms and Rule Evaluation Measures in Data Mining", *Fuzzy Sets and Systems*, 141(1) 59-88.

Nakashima, T., Nakai, G. and Ishibuchi, H. (2003). "Constructing Fuzzy Ensembles for Pattern Classification Problems", *Systems, Man and Cybernetics, IEEE International Conference on*, 4 3200-3205.

Quinlan, J. R. (1986). "Induction of Decision Trees", *MachineLearning*, 81-106.