

Clustering of Spatially Relevant Audio Data using Mel-Frequency Cepstrum for Diagnosis of Concrete Structure by Hammering Test

Jun Younes Louhi Kasahara¹, Hiromitsu Fujii¹, Atsushi Yamashita¹ and Hajime Asama¹

Abstract—Hammering test is a popular non-destructive testing method which automation is highly demanded for efficient diagnosis of concrete structures. The objective is to correctly determine if a hammering sound originated from a defect in the structure or not. In this paper, we present an unsupervised learning approach to automation of hammering test for diagnosis of concrete structures among others. Sound samples are clustered using fuzzy clustering while incorporating physical spatial information and Mel-Frequency Cepstrum is used in order to reproduce human hearing when conducting hammering test. Experiments using concrete test blocks showed good results, both in single and multiple defects cases.

I. INTRODUCTION

Concrete is extremely common in modern societies, especially in social infrastructures such as tunnels. In some cases, ageing and damages may lead them to structural failure [1]. In order to guarantee their safe use, careful maintenance is needed. Among all the operations taken to maintain these structures, the diagnosis for defects is critical since it is a decision-making step.

One popular diagnosis method is hammering: in this process, an operator hits perpendicularly the surface of the structure with a hammer and assesses the presence of defects from the perceived sound, as illustrated in Fig. 1. It is popular for being non-destructive and also not requiring heavy and/or precise equipment. However, this method requires a skilled operator to be able to correctly analyse the sound and given the huge population of structures in need of examination currently in service [2], testing them all with this traditional method reveals to be problematic. Therefore, the automation of the hammering test is highly demanded.

Several attempts to adapt the hammering test in an automatic form have been made: [3], [4], [5], [6] used supervised learning to correctly distinguish defect sounds from non-defect sounds. These approaches have given promising results, however their main drawback is the need of training sets: depending on various factors, concrete can greatly differ from one structure to another, even if they were made from the same batch, thus choosing the adequate training set is difficult.

In our previous method [7], based on clustering algorithm K-means++, the centroid of the cluster occupying the most of the tested structure's surface was used as a model for the non-defect sample to conduct diagnosis. However, this



Fig. 1: Hammering test conducted by a professional.

approach required a tedious threshold selection process by the user, limiting practicability. Our proposed method takes a new approach to this task, reinforcing clustering quality using spatial autocorrelation and Mel-Frequency Cepstrum (MFC), providing a robust and high-performing unsupervised learning method for discrimination of multiple defects.

II. CONCEPT

When witnessing human operators conducting hammering, it appeared clear that a hammering sample was not evaluated alone but relatively to other samples physically located around it: operators usually strike repeatedly following a line. This consideration motivated us to include the concept of spatial autocorrelation in our method, inspired by its application in computer vision [8]. This idea seems furthermore relevant since defects are localized and compact on a concrete structure.

Human operators are able to differentiate with relatively high accuracy defect and non-defect hammering samples. Understanding the exact physical phenomena that generate hammering samples would be complex, however the success of hammering as a testing method should indicate that the audio features characterizing hammering samples are within the human ear's perception range: humans are able to discriminate defect hammering sounds. MFC is a feature vector for audio data devised to simulate human hearing, widely used in speech recognition [9], [10] and gaining popularity in other related fields such as music information retrieval [11].

Another assumption is that most of the tested structure is non-defective. This means that it can be assumed that most of the hammering samples collected on a structure are non-defects. The justification for this is that for a heavily

¹J. Y. Louhi Kasahara, H. Fujii, A. Yamashita and H. Asama are with the Department of Precision Engineering, Faculty of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan. {louhi, fujii, yamashita, asama}@robot.t.u-tokyo.ac.jp

damaged structure, a simple visual inspection is enough to conduct diagnosis, defects in such high presence would appear on the surface and be blatant. Thus, this provides us with a statistical characterization for non-defect hammering samples.

III. METHOD

A. Feature vector

Considering a dataset $D = \mathcal{X}_1, \dots, \mathcal{X}_N$ composed of N hammering samples, a hammering sample $\mathcal{X}_i = \{\mathbf{l}_i, \mathbf{x}_i\}$ is defined as follows:

1) *Spatial information*: A physical location, point of contact between the hammer head and the structure's surface, noted $\mathbf{l}_i = (l_i^x, l_i^y)$ (assuming the tested surface is a plane).

The Euclidian distance is used to quantify physical distance between two hammering samples on the structure's surface, noted $\|\mathbf{l}_i - \mathbf{l}_j\|$.

2) *Audio information using MFC*: Mel-Frequency Cepstral Coefficients (MFCCs) of the sound resulting from the impact of the hammer head to the structure's surface, noted \mathbf{x}_i .

A metric based on the sample Pearson correlation coefficient, as defined in [7], was used to quantify similarity between MFCCs of two hammering samples. Given the MFCCs $\mathbf{a} = (a_l)_{l=1, \dots, M}$ and $\mathbf{b} = (b_l)_{l=1, \dots, M}$, the distance between them $d(\mathbf{a}, \mathbf{b})$ is computed as in eq. (1) and eq. (2), with M being the number of filters used during generation of MFCCs and \bar{a} and \bar{b} being the average coefficient of \mathbf{a} and \mathbf{b} respectively. The sample Pearson correlation coefficient has the advantage of including a zero mean and unit variance, enabling audio comparison of hammering samples despite not monitoring the input i.e. the impact force of the hammer head in our case.

$$r_{ab} = \frac{\sum_{l=1}^M [(a_l - \bar{a})(b_l - \bar{b})]}{\sqrt{\sum_{l=1}^M (a_l - \bar{a})^2} \sqrt{\sum_{l=1}^M (b_l - \bar{b})^2}} \quad (1)$$

$$d(\mathbf{a}, \mathbf{b}) = \frac{1 - r_{ab}}{2} \quad (2)$$

B. Fuzzy Clustering with Spatial Autocorrelation and MFC

1) Fuzzy C-Means Clustering (FCM) initialization:

For each sample is defined a set of fuzzy coefficients (u_{i1}, \dots, u_{iK}) , randomly initialized at the beginning, with K the number of clusters. Fuzzy coefficient u_{ij} is defined as quantification of hammering sample \mathcal{X}_i belonging to the j^{th} cluster [8].

2) *FCM update in MFC space*: At each iteration of FCM, the fuzzy coefficients are updated following eq. (3), with $\mathbf{c}_1, \dots, \mathbf{c}_K$ being the centroids of clusters with number of clusters K an user input and m a parameter controlling the fuzziness of the system.

$$u_{ij} = \frac{1}{\sum_{l=1}^N \frac{d(\mathbf{x}_l, \mathbf{c}_i)^{2/(m-1)}}{d(\mathbf{x}_j, \mathbf{c}_j)}} \quad (3)$$

3) *FCM update with Spatial information*: The concept of spatial autocorrelation can be defined as follows: samples located close to each another are likely to belong to the same cluster. In our particular case, this concept can be interpreted simply as: two hammering samples physically close on the structure's surface are likely to belong to the same cluster.

For each hammering sample \mathcal{X}_i , a neighbourhood $NB(\mathcal{X}_i)$ is defined as samples located in a disc of radius γ centred on the sample's physical position \mathbf{l}_i , as in eq. (4). Based on this neighbourhood, an estimator h_{ij} is used to calculate an average of the fuzzy coefficients in the neighbourhood, as in eq. (5), with $|NB(\mathcal{X}_i)|$ being the cardinality of the neighbourhood. This can be interpreted as the "expected" fuzzy coefficients for sample \mathcal{X}_i from other samples that are physically located near it: the average of $(u_{kj})_{k \in NB(\mathcal{X}_i)}$ translates the principle of spatial autocorrelation for the j^{th} cluster.

$$NB(\mathcal{X}_i) = (\mathcal{X}_k \in D \mid \|\mathbf{l}_i - \mathbf{l}_k\| \leq \gamma) \quad (4)$$

$$h_{ij} = \frac{1}{|NB(\mathcal{X}_i)|} \sum_{k \in NB(\mathcal{X}_i)} u_{kj} \quad (5)$$

At each iteration of FCM, along with the update related to the clustering in MFC space, a "smoothing" is effected as in eq. (6) in order to take into account the localization of each hammering sample, with p and q weighting exponents on each fuzzy membership, regulating the contributions of respective sources.

$$u_{ij} \rightarrow \frac{u_{ij}^p h_{ij}^q}{\sum_k u_{kj}^p h_{kj}^q} \quad (6)$$

4) *Defuzzification*: The obtained fuzzy clustering is finally converted to a crisp clustering by maximum membership: each hammering sample is assigned to the cluster that has the highest fuzzy coefficient for that hammering sample as in eq. (7), with n_i being the cluster attributed to hammering sample.

$$n_i = \arg \max_j (u_{ij}) \quad (7)$$

C. Cluster identification

As stated earlier, the cluster occupying most of the tested structure's surface is identified as the non-defect cluster. In order to evaluate the surface occupied by each cluster, a weight w_i is devised for each hammering sample based on its physical distance to its nearest neighbour r_i , as shown in eq. (8) and eq. (9): this weight illustrates the area where that hammering sample is relevant i.e. representative of.

$$r_i = \frac{1}{2} * \min_{\forall j \in D} \|\mathbf{l}_i - \mathbf{l}_j\| \quad (8)$$

$$w_i = \pi * r_i^2 \quad (9)$$

From there, weights associated for each centroid w_{c_j} can be calculated as in eq. (10) by summing the weights of hammering samples belonging to its cluster.

$$w_{\mathbf{c}_k} = \sum_{\mathcal{X}_i \in \mathbf{c}_j} w_i \quad (10)$$

Finally, the centroid with the highest weight is identified as the centroid of the non-defect cluster $\mathbf{c}_{\text{non-defect}}$, as in eq. (11).

$$\mathbf{c}_{\text{non-defect}} = \mathbf{c}_{\arg \max_k (w_{\mathbf{c}_k})} \quad (11)$$

A pseudo-algorithm briefly describing our proposed method is shown in Algorithm (1).

Data: dataset of hammering samples D , number of clusters K

Initialization;

foreach hammering sample $\mathcal{X}_i \in D$ **do**
| assign random fuzzy coefficients u_{ij} ;

end

Loop;

while termination criterion is not met **do**

foreach cluster centroid \mathbf{c}_j **do**

 conduct centroid update;

$$\mathbf{c}_j = \frac{\sum_i u_{ij}^m \mathbf{x}_i}{\sum_i u_{ij}^m}$$

end

foreach hammering sample $\mathcal{X}_i \in D$ **do**

 conduct regular FCM update;

$$u_{ij} = \frac{1}{\sum_{i=1}^N \frac{d(\mathbf{x}_i, \mathbf{c}_i)^{2/(m-1)}}{d(\mathbf{x}_i, \mathbf{c}_j)^{2/(m-1)}}};$$

 build spatial estimator;

$$h_{ij} = \frac{1}{|NB(\mathcal{X}_i)|} \sum_{k \in NB(\mathcal{X}_i)} u_{kj};$$

 balance results;

$$u_{ij} \rightarrow \frac{u_{ij}^p h_{ij}^q}{\sum_k u_{kj}^p h_{kj}^q}$$

end

end

Defuzzification;

foreach hammering sample $\mathcal{X}_i \in D$ **do**
| assign sample to cluster with highest fuzzy coefficient;

$$n_i = \arg \max_j (u_{ij})$$

end

Identification;

foreach hammering sample $\mathcal{X}_i \in D$ **do**

 find distance to nearest hammering sample ;

$$r_i = \frac{1}{2} * \min_{j \in D} \|\mathbf{l}_i - \mathbf{l}_j\|;$$

 compute weight;

$$w_i = \pi * r_i^2;$$

end

foreach cluster centroid \mathbf{c}_j **do**

 sum member sample weights;

$$w_{\mathbf{c}_j} = \sum_{\mathcal{X}_i \in \mathbf{c}_j} w_i$$

end

cluster with highest weight is the non-defect cluster;

$$\mathbf{c}_{\text{non-defect}} = \mathbf{c}_{\arg \max_k (w_{\mathbf{c}_k})}$$

Algorithm 1: Pseudo algorithm of the proposed method.

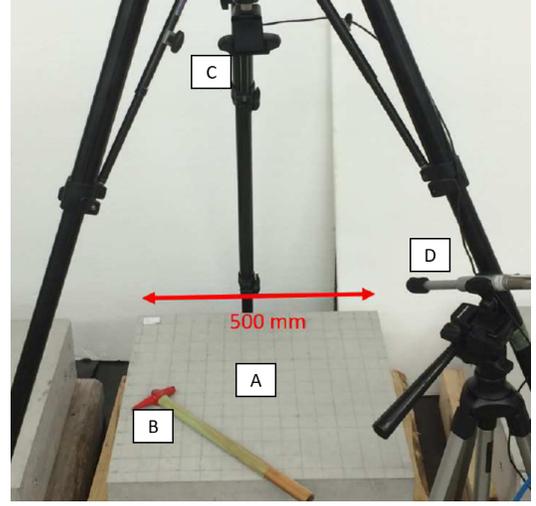


Fig. 2: Experimental environment showing the concrete block (A), hammer (B), camera (C) and microphone (D).

IV. EXPERIMENTAL SETUP

Several concrete test blocks containing man-made defects simulating natural ones were made. For practical reasons, each block contained a single defect so that the defect area would be significant enough to be relevant. The experimental environment is shown in Fig. 2: blocks were hit at several locations to test our proposed method's ability to robustly discriminate defect hammering samples.

In our case, the location of each hammering sample was a simple 2D position, collected by tracking a red-painted hammer head using computer vision. The used hammer was a KTC UDHT-2 (length 380 mm, weight 160 g, head diameter 16 mm), commonly used in hammering test by professionals.

Sound was recorded at 44.1 kHz using a Behringer ECM8000 condenser microphone coupled with a Roland UA25EX sound board and a laptop PC for data analysis. A simple trigger was implemented to conduct clipping to get each hammering sound as a single sample. Audio data were first transformed into Fourier spectrum by Fast Fourier Transform (FFT) using a window size of 1024 and then MFCCs were computed using 26 triangular filters in the 300-8000 Hz frequency band.

Number of clusters K was set to 2, p and q were set to unity and γ was set so that every sample would have at least one neighbour. Termination criterion was 100 iterations.

A. Single defect case

The concrete block used for the purpose of testing detection of a single defect is shown in Fig. 4(a). Referring to the schematic presented in Fig. 3, $\alpha=30$ deg., $l=200$ mm, $d=115.5$ mm and $L=230.9$ mm. 462 hammering samples were collected on this block, with a human operator conducting the hammering action: 272 non-defect and 190 defect samples.

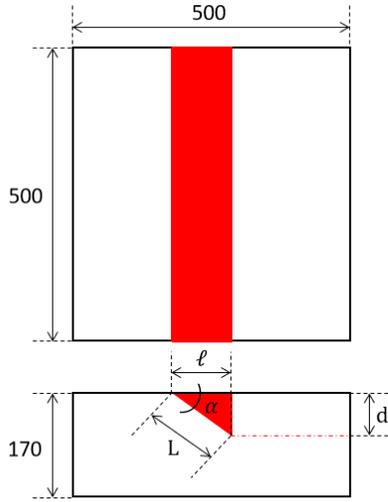


Fig. 3: Generic schematic of the tested concrete blocks. Defect area shown in red, dimensions in mm.

B. Multiple defect case

In order to test the multiple defect scenario, two concrete blocks each containing a man-made delamination were put together. Referring to the schematic presented in Fig. 3 and to Fig. 5(a), for the block on the left, $\alpha=15$ deg., $d=40$ mm, $l=149.3$ mm and $L=154.5$ mm. For the block on the right, $\alpha=15$ deg., $l=200$ mm, $d=53.6$ mm and $L=207.1$ mm. 270 hammering samples were collected in total on these blocks, with a human operator conducting the hammering action: 155 non-defect and 115 defect samples.

V. RESULTS

In order to quantify the classification performance, from the amount of true positives TP (defect samples detected as defects), true negatives TN (non-defect samples detected as non-defects), false positives FP (non-defect samples detected as defects) and false negatives FN (defect samples detected as non-defects), the values of precision, recall and accuracy that are normally used as performance indices of classification were computed following eq. (12), eq. (13) and eq. (14).

$$\text{precision} = \frac{TP}{TP + FP} \quad (12)$$

$$\text{recall} = \frac{TP}{TP + FN} \quad (13)$$

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (14)$$

Our previous method [7], FCM using Fourier spectrum, FCM with spatial information using Fourier spectrum, FCM using MFC and our proposed method (FCM with spatial information using MFC) were tested so that the respective influence of spatial information and MFC on the classification result could be clearly distinguished. Classification results obtained in both the single defect and multiple defect case are shown in Table I, with values of precision, recall and accuracy in the form of percentage for better readability.

A. Single defect case

As shown in Fig. 4(b), our previous method, with threshold value for detection manually selected in order to obtain the best output, failed in this particular case of single defect detection: almost no sample inside the red dashed frame, indicating the defect area, has been classified as defect and most of the left-side non-defect portion of the concrete block was detected as defect. This could be explained by the high defect to non-defect area ratio compared to previous concrete blocks used in [7]: this would influence badly the model generation and selection step.

If using Fourier spectrum, both FCM and FCM with spatial information, Fig. 4(c) and Fig. 4(d), failed in the cluster identification step too, although the clustering itself, separating samples along the delimitation, at the left side of the defect area, does not seem totally irrelevant. In this case, the addition of spatial information slightly influenced the output by removing the small patches of samples wrongly identified as defects on the right-side non-defect area.

Introduction of MFC significantly improved results for FCM, as shown in Fig. 4(e): almost all samples inside the red dashed frame and only a few outside it have been classified as defects.

Adding spatial information (Fig. 4(f)) enabled to further refine the result by removing the patch of mislabelled non-defect samples outside the red dashed frame and achieving high values of precision, recall and accuracy (100%, 98.95% and 99.57% respectively) as shown in Table I.

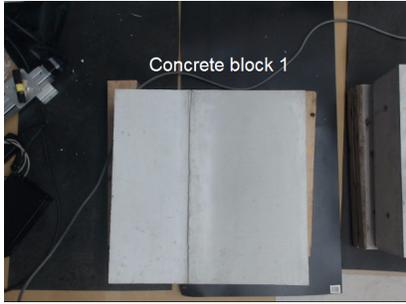
B. Multiple defect case

As shown in Fig. 5(b), our previous method with manual threshold exhibited a better result than in the single defect case: most samples inside the right red dashed frame have been correctly detected as defects as well as about half the samples in the left red dashed frame, with few non-defects classified as defects.

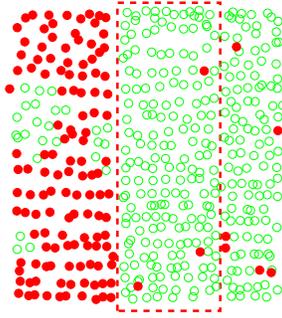
If using Fourier spectrum, both FCM and FCM with spatial information, Fig. 5(c) and Fig. 5(d), showed irregular quality outputs. This is most certainly due to the initialization being random as well as the dataset's composition: initial conditions heavily influenced the output i.e. clusters were not well separated. This issue was not apparent in the single defect case where outputs of algorithms stayed constant over 100 runs.

Introduction of MFC again significantly improved results for FCM, as shown in Fig. 5(e) and stabilized the output: only a few samples have been mislabelled as defects outside the frames and inside them, most samples have been detected as defects. This would indicate that indeed, MFC space contains well separated clusters, enabling FCM to return stable and correct results even with random initialization [12].

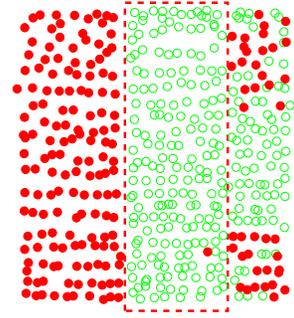
Adding spatial information (Fig. 5(f)) gave an excellent result, with a precision of 100%: samples labelled as defects were all inside the defect area: the few isolated samples wrongly detected as defects were removed and the detection rate inside the defect area, especially on the edges where delaminations were deep and made detection harder, increased.



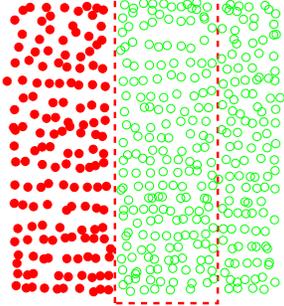
(a) Picture of tested concrete block



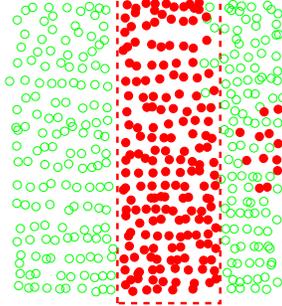
(b) Previous method



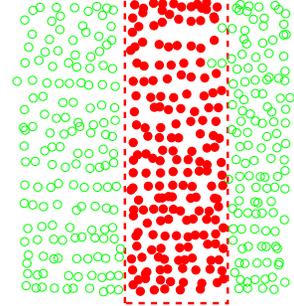
(c) FCM with Fourier Spectrum



(d) Spatial FCM with Fourier Spectrum

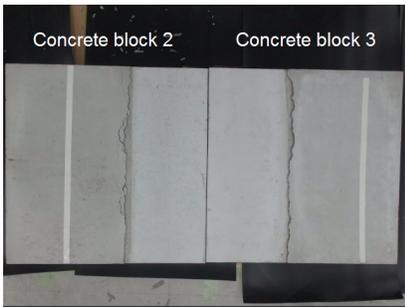


(e) FCM with MFC

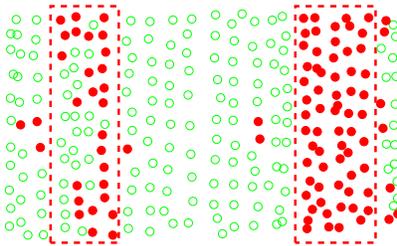


(f) Spatial FCM with MFC

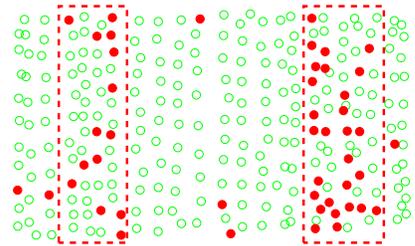
Fig. 4: Results for the single defect case, each node represents a hammering sample: green circles are samples classified as non-defects, red dots are samples classified as defects. Red dashed frames show defect area (ground truth).



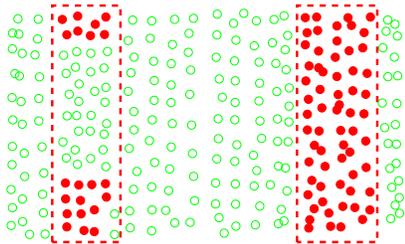
(a) Picture of tested concrete block



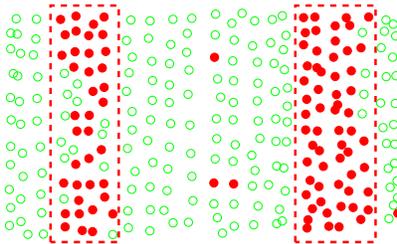
(b) Previous method



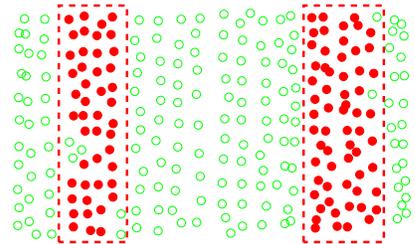
(c) FCM with Fourier Spectrum



(d) Spatial FCM with Fourier Spectrum



(e) FCM with MFC



(f) Spatial FCM with MFC

Fig. 5: Results for the multiple defect case, each node represents a hammering sample: green circles are samples classified as non-defects, red dots are samples classified as defects. Red dashed frames show defect area (ground truth).

TABLE I: Performance of various methods in the single and multiple defect cases. Average, best (+) and worst (-) results over 100 runs of the method when applicable are shown. Letters in the first column indicate correspondences with results shown in Fig. 4 and Fig. 5.

Method	Single defect (Fig. 4)			Multiple defect (Fig. 5)		
	Precision	Recall	Accuracy	Precision	Recall	Accuracy
Previous (b)	1.27%	0.5%	42.3%	86.9%	80.2%	86.4%
FCM with Fourier Spectrum (c)	0.5%	0.5%	17.7%	18.9%	12.6%	59.3%
	(+) 0.5%	(+) 0.5%	(+) 17.7%	(+) 100%	(+) 99.1%	(+) 91.2%
Spatial FCM with Fourier Spectrum (d)	(-) 0.5%	(-) 0.5%	(-) 17.7%	(-) 0%	(-) 0%	(-) 22.1%
	0%	0%	26.4%	97.6%	58.5%	81.6%
FCM with MFC (e)	(+) 0%	(+) 0%	(+) 26.4%	(+) 100%	(+) 95.7%	(+) 98.2%
	(-) 0%	(-) 0%	(-) 26.4%	(-) 51.6%	(-) 55.2%	(-) 58.8%
Proposed (Spatial FCM with MFC) (f)	93.9%	96.8%	96.1%	96.2%	86.2%	92.7%
	(+) 93.9%	(+) 96.8%	(+) 96.1%	(+) 96.2%	(+) 86.2%	(+) 92.7%
Previous (b)	(-) 93.9%	(-) 96.8%	(-) 96.1%	(-) 96.2%	(-) 86.2%	(-) 92.7%
	100%	99.0%	99.6%	100%	88.8%	95.2%
Proposed (Spatial FCM with MFC) (f)	(+) 100%	(+) 99.0%	(+) 99.6%	(+) 100%	(+) 88.8%	(+) 95.2%
	(-) 100%	(-) 99.0%	(-) 99.6%	(-) 100%	(-) 88.8%	(-) 95.2%

Again, compared to the results using our previous method, with precision, recall and accuracy at 86.9%, 80.2% and 86.4% respectively, the improvement is noticeable: precision at 100%, recall at 88.8% and accuracy at 95.2%, as shown in Table I.

VI. CONCLUSIONS

The method described in this paper was able to exhibit good performance in the presence of single and multiple delaminations, outperforming our previous work in both cases by replacing analysis in Fourier spectrum space with MFC space and using the concept of spatial autocorrelation. To this regard, introduction of MFC ultimately showed that a definitive feature differentiating defect and non-defect samples was caught, at least partly, in this feature and that the addition of spatial information in the analysis helped in the generation of clusters that match defect and non-defect areas.

In this present paper, the tested multiple defect case was composed of only delaminations and therefore the setting $K = 2$ yielded satisfying results. If the defects were of different nature, it can be expected that another setting for K would be more appropriate, and in the field it can be expected to find several types of defects i.e. defects other than delaminations. Therefore, in future works, we would like to produce automatic cluster number selection along with the associated seeding procedure in order to tackle the case of multiple defects of different nature. Furthermore, elaboration of a feature specific to this task could further refine analysis. Finally, tests on real, natural defects on concrete structures in the field are also to be conducted.

ACKNOWLEDGMENT

This work was supported in part by Council for Science, Technology and Innovation, "Cross-ministerial Strategic Innovation Promotion Program (SIP), Infrastructure Maintenance, Renovation, and Management" (funding agency: NEDO), JSPS KAKENHI Grant Number JP16H06680, and Institute of Technology, Tokyū Construction Co., Ltd.

REFERENCES

- [1] Yasuhiro Mori and Bruce R. Ellingwood. Reliability-based service-life assessment of aging concrete structures. In *Journal of Structural Engineering*, volume 119, pages 1600–1621, 1993.
- [2] Ryoki Takada and Noboru Oishi. Priority issues of infrastructure inspection and maintenance robot: A part of COCN 2012 project disaster response robot and its operational system. In *Humanitarian Technology Conference (R10-HTC), 2013 IEEE Region 10*, pages 166–171. IEEE, 2013.
- [3] Gang Zhang, Ronald S Harichandran, and Pradeep Ramuhalli. Application of noise cancelling and damage detection algorithms in nde of concrete bridge decks using impact signals. In *Journal of Nondestructive Evaluation*, volume 30, pages 259–272. Springer, 2011.
- [4] Gang Zhang, Ronald S. Harichandran, and Pradeep Ramuhalli. An automatic impact-based delamination detection system for concrete bridge decks. In *NDT & E International*, volume 45, pages 120–127. Elsevier, 2012.
- [5] Hiromitsu Fujii, Atsushi Yamashita, and Hajime Asama. Defect detection with estimation of material condition using ensemble learning for hammering test. In *Robotics and Automation (ICRA), 2016 IEEE International Conference on*, pages 3847–3854. IEEE, 2016.
- [6] Ye Jiaying, Masaya Iwata, Kobayashi Takumi, Masahiro Murakawa, Higuchi Tetsuya, Yuichi Kubota, Toshiya Yui, and Kiyoshi Mori. Statistical impact-echo analysis based on Grassmann manifold learning: Its preliminary results for concrete condition assessment. In *EWSSH-7th European Workshop on Structural Health Monitoring*, pages 1349–1356, 2014.
- [7] Jun Younes Louhi Kasahara, Hiromitsu Fujii, Atsushi Yamashita, and Hajime Asama. Unsupervised learning approach to detection of void-type defects in concrete structure using hammering and clustering. In *The 6th International Conference on Advanced Mechatronics*, number 6, pages 319–320, 2015.
- [8] Keh-Shih Chuang, Hong-Long Tzeng, Sharon Chen, Jay Wu, and Tzong-Jer Chen. Fuzzy c-means clustering with spatial information for image segmentation. In *Computerized Medical Imaging and Graphics*, volume 30, pages 9–15. Elsevier, 2006.
- [9] Todor Ganchev, Nikos Fakotakis, and George Kokkinakis. Comparative evaluation of various MFCC implementations on the speaker verification task. In *Proceedings of the SPECOM*, volume 1, pages 191–194, 2005.
- [10] Beth Logan et al. Mel frequency cepstral coefficients for music modeling. In *International Society for Music Information Retrieval*, 2000.
- [11] Meinard Müller. Information retrieval for music and motion. Springer, 2007.
- [12] Dzong L Pham and Jerry L Prince. Adaptive fuzzy segmentation of magnetic resonance images. In *IEEE transactions on medical imaging*, volume 18, pages 737–752. IEEE, 1999.