Substituting Spatial by Temporal Information in Clustering of Audio Data for Defect Diagnosis

The University of Tokyo ○ Louhi Kasahara Jun Younes, Miyagusuku Renato, Fujii Hiromitsu
Yamashita Atsushi, Asama Hajime

Concrete is an omnipresent material in social infrastructures, requiring careful diagnosis for defects to ensure users' safety. The hammering test is a traditional non-destructive testing method based on sound to find defects non-visible from the surface of structures. Our previous works successfully demonstrated the possibility of discriminating defect hammering samples using fuzzy c-means with spatial information. However, spatial information can sometimes be difficult to obtain in practice. The present paper investigates the possibility to substitute spatial information by temporal information. Early results of experiments using concrete test blocks indicate that although performance is degraded, better results are obtained than using only audio information.

1. Introduction

Concrete is omnipresent in modern societies, especially in social infrastructures such as tunnels and bridges. Deterioration by ageing or environmental factors can lead to severe failure of those structures, endangering users [1]. Diagnosis for defects is therefore of critical importance. The hammering test is a method used in regular inspections: a human hits the surface of the structure with an inspection hammer, and the returned sound is used to assess if the hit spot is a defect or not. This method requires a skilled operator to correctly interpret hammering sounds and the growing number of structures to test demands a more efficient alternative [2]. Therefore, automation of the hammering test is highly desirable.

Related studies are mainly based on machine learning, consisting of supervised and unsupervised approaches. In [3] a modified Independent Component Analysis was used along a Radial Basis Function neural network based on Mel-Frequency Cepstrum Coefficients (MFCCs) to classify hammering samples. In [4] the authors used ensemble learning techniques with time-frequency analysis in order to achieve a classifier able to detect defects and classify them according to their depth from the surface. These approaches have yielded remarkable results. However, in this particular case training sets are difficult to obtain. Furthermore, concrete is highly variable, i.e. each structure has distinct sounds.

Unsupervised learning approaches have been proposed in our previous works: in [5], spatial Fuzzy C-Means (FCM) and Mel-Frequency Cepstrum Coefficients (MFCCs) were employed and the proposed methodology, illustrated in Fig. **1**, showed high performance in the case of single and multiple delaminations. Addition of spatial information is quite effective, however it can sometimes be difficult to obtain in actual inspection sites.

Therefore, in this paper we investigate the possibility of substituting spatial information by temporal information, i.e. replacing the position of hammering samples by the data collection order.

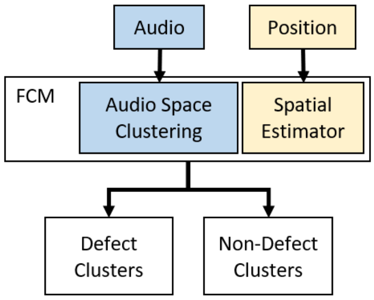


Fig. 1: FCM with Spatial estimator for defect diagnosis.

2. Method

Spatial FCM is a well known fuzzy clustering algorithm in fields such as computer vision where it was notably used for image segmentation [6]. The core idea is to take advantage of the fuzzy aspect to inject spatial information in the clustering process. This algorithm was successfully modified and used for defect hammering sample detection in [5]. Given a hammering sample $\mathscr{X}_i$ with its position $\mathbf{l}_i$, a spatial neighbourhood $NB(\mathscr{X}_i)$ was defined, as in eq. (1): a disc of radius γ based on Euclidian distance. Based on this neighbourhood, a spatial estimator h_{ij} , was defined as the average of corresponding fuzzy memberships, as shown in eq. (2), with $|NB(\mathscr{X}_i)|$ being the number of neighbours for sample $\mathscr{X}_i$.

$$NB(\mathscr{X}_i) = (\mathscr{X}_i \in D \ \mid \ \left\|\mathbf{l}_i - \mathbf{l}_j\right\| \leq \gamma) \tag{1}$$

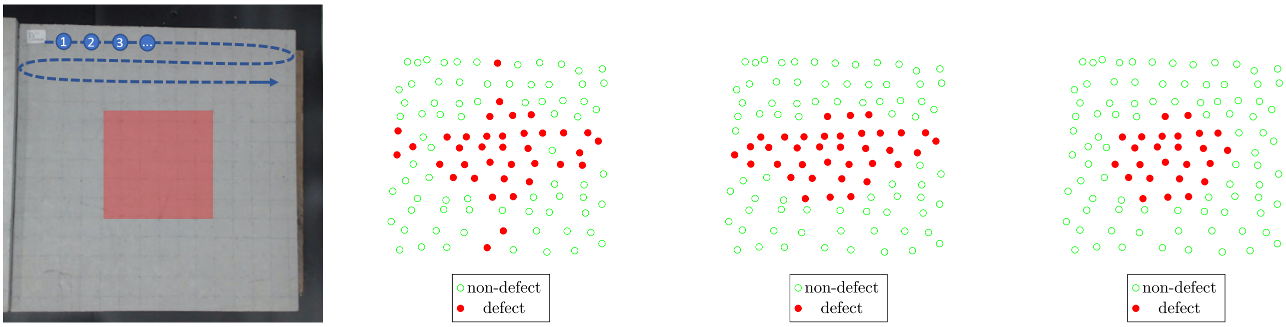
$$h_{ij} = \frac{1}{|NB(\mathscr{X}_i)|} \sum_{k \in NB(\mathscr{X}_i)} u_{kj} \tag{2}$$

If we suppose that the data sampling was conducted roughly following a grid and that the dataset $D = \mathscr{X}_1, ..., \mathscr{X}_N$ is ordered by data collection time, i.e. if $i < j$ then $\mathscr{X}_i$ was collected before $\mathscr{X}_j$, a substitute for spatial estimator can be defined as in eq. (3) (the first and last sample in D are not affected):

$$\forall i \in [2, ..., N-1] \quad h_{ij} = \frac{1}{2}(u_{(i-1)j} + u_{(i+1)j}) \tag{3}$$

3. Experiments

Regular FCM, FCM with temporal estimator and FCM with spatial estimator were tested on a concrete test block containing a polystyrene cuboid at a depth of 30mm to simulate a defect, as shown in Fig. **2**(a). The dataset was composed of 111 samples:



(a) Concrete test block used in experiment (red area shows defect, data collection order shown with blue arrow).

(b) Output using regular FCM.

(c) Output using FCM + temporal.

(d) Output using FCM + spatial.

Fig. 2: Diagnosis outputs using regular FCM, FCM with temporal estimator and FCM with spatial estimator.

Table 1: Diagnosis performance for each method.

	Precision	Recall	Accuracy
Regular FCM (b)	62.2%	88.5%	84.7%
FCM + temporal (c)	66.7%	92.3%	87.4%
FCM + spatial (d)	92.0%	88.5%	95.5%

85 non-defects and 26 defects. Audio data was inputted into the algorithms as MFCCs, a sound feature vector popular in speech recognition, truncated to 10 coefficients. Location position of hammering samples was obtained by color tracking the hammer head. Hammering samples were collected roughly following horizontal lines, beginning at the upper left portion of the block.

4. Results and Discussion

The diagnosis output of each method are shown in Fig. **2**(b), **2**(c) and **2**(d): red dots represent samples classified as defect and green circle samples classified as non-defect. In Table **1**, the classic performance values of *precision*, *recall* and *accuracy* are shown.

Regular FCM has issues with outliers and samples located at the edges of the test block, where border conditions, and thus returned sound, tend to change slightly. FCM with temporal estimator returned an overall better result, outperforming regular FCM in all three performance indicators in this particular case. However, it can be seen that the output is heavily influenced by the data collection order: since hammering samples were collected horizontally from top to bottom, the defect area is "stretched" sideways. FCM with spatial estimator on the other hand returned an excellent result, closely matching the test block's structure. Precison and accuracy are the highest among the three tested method, at 92.0% and 95.5%, respectively. Recall is slightly lower than FCM with temporal estimator, at 88.5%, but this should be nuanced with the rather high number of false positives returned by the latter.

Globally, replacing spatial estimator by temporal estimator degraded diagnosis performance. Since less information is used, this was expected. However, it is interesting to note that the returned performance still outperforms regular FCM. Even temporal information can aid in the diagnosis for defect hammering samples, additional information helps in the diagnosis.

5. Conclusion

In this paper, we investigated the possibility of substituting spatial information by temporal information for clustering of audio data for detection of defects in concrete structures. Preliminary results hint that temporal information can make up for spatial information up to a certain point. However, its sensitivity to the data collection pattern, i.e. correlation between temporal and spatial information, should be carefully taken into account.

In future works, we would like to clarify the influence of hammering pattern on the output when using temporal information. Also, we would like to establish a temporal estimator more robust against hammering pattern and try to match more closely the detection performance with spatial information.

Acknowledgments

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, Tokyu Construction Co., Ltd.

References

1) Y. Mori and B. R. Ellingwood, "Reliability-based service-life assessment of aging concrete structures," *Journal of Structural Engineering*, vol. 119, pp. 1600–1621, 1993.
2) R. Takada and N. 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*, pp. 166–171, IEEE, 2013.
3) G. Zhang, R. S. Harichandran, and P. Ramuhalli, "An automatic impact-based delamination detection system for concrete bridge decks," *NDT & E International*, vol. 45, no. 1, pp. 120–127, 2012.
4) H. Fujii, A. Yamashita, and H. Asama, "Defect detection with estimation of material condition using ensemble learning for hammering test," in *Proceedings of the International Conference on Robotics and Automation (ICRA)*, pp. 3847–3854, IEEE, 2016.
5) J. Y. Louhi Kasahara, H. Fujii, A. Yamashita, and H. Asama, "Clustering of spatially relevant audio data using Mel-frequency cepstrum for diagnosis of concrete structure by hammering test," in *Proceedings of the International Symposium on System Integration*, IEEE/SICE, 2017.
6) K.-S. Chuang, H.-L. Tzeng, S. Chen, J. Wu, and T.-J. Chen, "Fuzzy c-means clustering with spatial information for image segmentation," *Computerized Medical Imaging and Graphics*, vol. 30, no. 1, pp. 9–15, 2006.