Proceedings of the 21st ICSMGE, Vienna, Austria, 14 – 19 June 2026. Pistrol, Adam & Schweiger (eds.)
Published by: ÖGG, Austrian Society for Geomechanics, Salzburg, Austria, ISBN 978-3-9503898-4-5
https://doi.org/10.53243/ICSMGE2026-440

Directional sensitivity of stereo microphones in cavity localization

Kingsley Asare Boakye, Changho Lee
Department of Architecture and Civil Engineering, Chonnam National University, Gwangju, South Korea,
changho@jnu.ac.kr

ABSTRACT: Air-coupled impact-echo has been widely used for the detection and identification of damage in structures. However, either the use of single microphone or MEMS in detecting cavities has many limitations such as low sensitivity and inability to filter high frequency response respectively, and it is pertinent to provide a robust method of clearly delineating abnormalities in structures. This study introduces stereo microphone system and averaging and convolution signal processing techniques of the stereo microphone system for void detection. In this present study, voids of varying widths were designed into a concrete structure, sound waves were generated by impacting an instrumented hammer and responses recorded with stereo microphones. Signal processing of averaging and convolution were then applied to the signals and analysis of the signals were based on the FFT to distinguish defective and sound areas of the structures. The study found that generally as the size of the void increased cavity detection increased with signal response from the stereo microphone showing the highest detection rate as compared to the mono signal response of 53%. Convoluted signal response had the lowest error rate of 25% clearly indicating the improvement of cavity detection using stereo microphones.

KEYWORDS: air-coupled impact-echo, stereo-microphone, FFT, alignment, signal processing, convolution, beamforming.

1 INTRODUCTION

The integrity of tunnels, bridges and in general civil structures is critical to the safety, durability and overall functionality of infrastructure. Defects such as voids, delamination, and cracks in these systems can lead to catastrophic failures if undetected. Over the past decade, the impact-echo (IE) method has emerged as a viable non-destructive testing (NDT) technique for identifying subsurface anomalies in concrete structures. By analyzing the frequency response of stress waves generated by a mechanical impact, IE enables the detection of defects such as interfacial debonding in CRTS III slab tracks (Jiang et al., 2021) and segment-grout voids in tunnels (Yao, Chen and Abula, 2018). Despite its effectiveness, conventional IE methods face challenges in noisy environments, limited spatial resolution, and ambiguity in distinguishing defect-induced echoes from reflections caused by reinforcing bars or geometric boundaries (Yeh, Liu and Hsu, 2018; Chang, Yu and Lin, 2019).

Recent advancements in signal processing, such as wavelet transforms (Yu, Lin and Chang, 2021),empirical mode decomposition (Yumnam, Ghosh and Gupta, 2023), and machine learning (Xu and Yu, 2021; Inaba, Tanigawa and Naito, 2023), have improved defect characterization by enhancing time-frequency analysis and automating signal interpretation. However, these approaches predominantly rely on single-channel sensor data, which restricts the ability to leverage spatial wave propagation patterns. Multi-sensor configurations, such as ultrasonic arrays (Kuchipudi et al., 2023) and distributed fibre optics (Gong et al., 2023), have demonstrated the value of spatial data fusion in improving defect localization and quantification. Yet, such systems often require complex instrumentation or are impractical for rapid field inspections.

To address these limitations, this study proposes a dual microphone-based IE system integrated with advanced signal processing techniques specifically averaging and convolution. The stereo configuration allows for beamforming and directional filtering to suppress ambient noise while enhancing defect-related signals. Frequency domain analysis is used to evaluate the signals for localizing the cavity which differentiates the concrete and cavity regions.

2 EXPERIMENTAL STUDY

2.1 *Laboratory setup*

A 500 mm x 500 mm x 250 mm concrete specimen with a simulated hollow was created and positioned in the bottom middle to identify the cavity in the concrete. The surface was marked with a grid of 50 x 50 mm^2 test sites, and an air-coupled impact-echo test was conducted at each grid location. An instrumented hammer with a 6.3 mm tip diameter was used to create sound waves, and a microphone was used as a receiver. The microphone utilized in this investigation measured up to 20 kHz in operating frequency, had a ½-inch diameter, and a sensitivity of 50 mV/Pa (378B02, PCB Piezotronics NY, USA). The microphone was installed on the surface of the concrete with a wave guide (Kang et al., 2021) of height 20 mm, acting as a noise reducer and placing the microphone at the desired height.

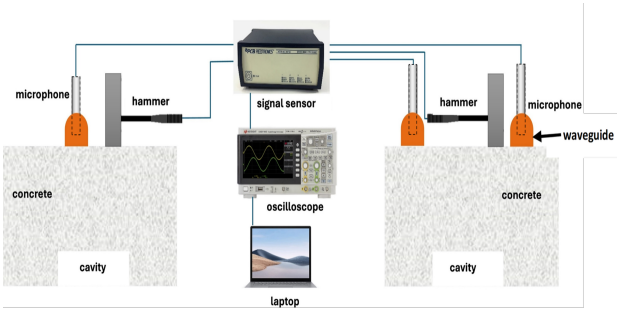


Figure 1. Experimental setup.

Two microphones with identical specifications as shown in Figure 1 were used for the impact echo test, and they were positioned 100 mm apart. The microphones create and record sound waves when the hammer strikes the concrete surface. These waves are then amplified and filtered by a sensor signal conditioner (PCB 482C64, PCB Piezotronics, NY, USA), converted to digital form on an oscilloscope (DSOX2014A, Keysight Technologies, CA, USA), and examined on a computer. 16660 sample points were used to sample the signals at a frequency of 167 kHz.

2.2 *Signal processing*

Several signals are collected from both the mono and stereo microphone systems. These signals are then detrended to

remove the effect of DC component on the signals and further remove noise from the signal, a 4th order Butterworth bandpass filter with high pass of 1.5kHz is applied and later normalized for uniformity.

Advanced processing of the signals received from the stereo-microphone system included explicitly these two major steps: averaging, and a convolution as shown in Figure 2. To enhance the clarity and reliability of stereo microphone data, two normalized input signals are combined into a single average signal hereby referred to as averaging. This mode captures shared characteristics of both signal responses while reducing noise and individual fluctuations (Hassan and Anwar, 2010).

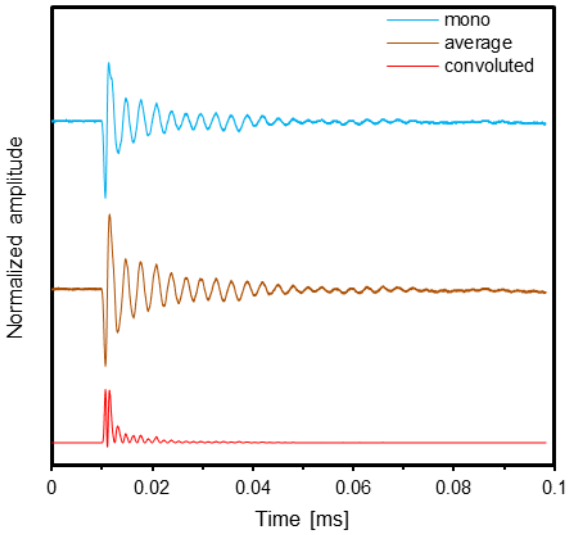


Figure 2. Response signals.

Convolution was applied by element-wise multiplying of the two (2) signals. This generates a signal that clearly reveals interactions of signals over a certain period (Akita, 2016). Fast Fourier Transform (FFT) is then used to analyze the signals to detect sound and defected areas in the concrete structure (Li et al., 2014; Yu, Lin and Chang, 2021) followed by a qualitative index k, which is represented in equation 1.

$$\left(1 - \left|\frac{A_O - A_E}{A_E}\right|\right) \times 100\% \quad (1)$$

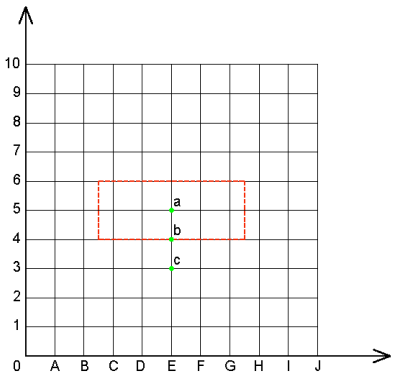
Where A_O represents observed defect area and A_E represents expected defect area.

3 RESULTS AND DISCUSSION

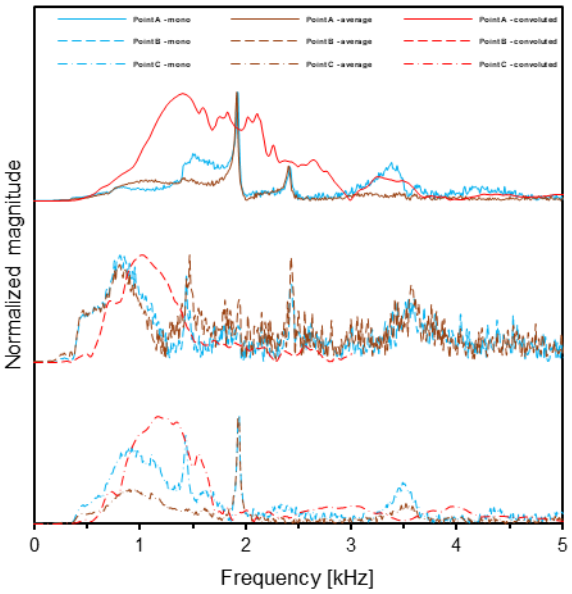
An examination of the spectral responses at three spatial points, Point A (defect area), Point B (edge of defect), and Point C (sound area), using three signal responses, mono, average, and convoluted, based on FFT analysis is discussed in this section in conjunction with detection rate and error analysis on cavity detection.

Figure 3 shows the frequency domain analysis of 100 mm width. As illustrated, there is no significant shift in frequency from Point A to Point C with the dominant spectral features, particularly around 2 kHz, remaining consistent across all points. At Point A, the mono signal response presents a moderate magnitude profile, gradually rising around 1.5 kHz and forming a broad peak between 2 and 3.5 kHz before tapering off. The average signal response closely follows this trend but appears smoother and more uniform, slightly reducing the fluctuations seen in the mono signal response. The convoluted signal response, in contrast, exhibits a significantly elevated magnitude and a wider spectrum of frequency ranging from 1 – 2.5 kHz.

Point B indicates that both the mono and average signal responses introduce more pronounced variations between 1 and 3.5 kHz, suggesting enhanced spectral detail. The convoluted signal response on the other hand displays a notably different profile, marked by well-defined peaks around 1 kHz. These features are sharper and more structured than those in the mono and average signals, indicating higher sensitivity in specific frequency bands.



(a)



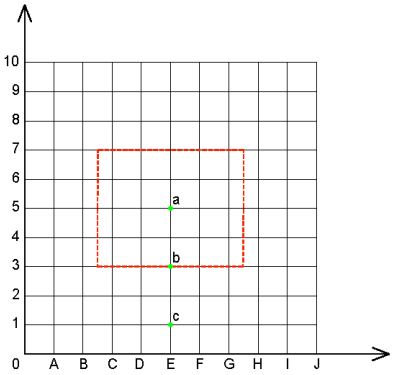
(b)

Figure 3. Frequency domain analysis of 100 mm width. (a) gridlines; (b) peak frequency of points.

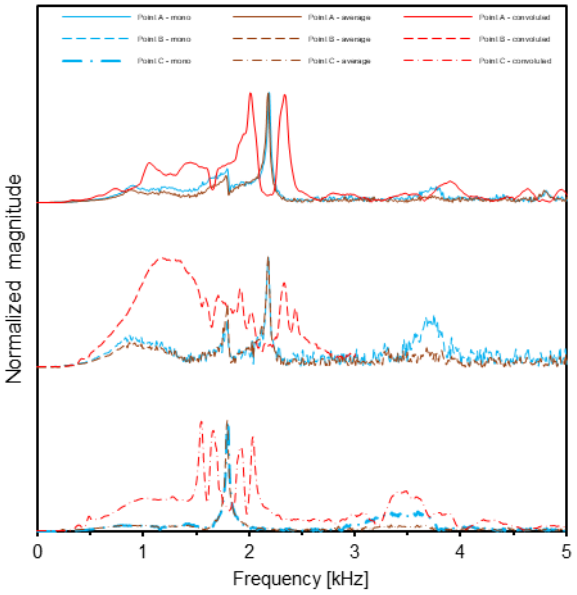
At point C, it is evident that the mono signal maintains a generally low magnitude with a gentle rise between 1 and 2 kHz and again at 3.5 kHz. The average signal shares this shape but appears slightly lower with more pronounced mid-frequency fluctuations. The convoluted signal reveals a sharper, more concentrated profile with a dominant peak near 2 kHz, followed by a rapid decline. Additional activity from 0.5 to 2.5 kHz gives the convoluted spectrum a more complex and focused appearance, in contrast to the broader, smoother responses of the other two signals.

In Figure 4, the primary frequency content remains located in the same general band, centered around 1.5 to 2.5 kHz. However, the intensity, sharpness, and definition of spectral peaks change progressively from the defect to the sound area.

For the convoluted signals, Point A exhibits sharply defined, high-magnitude peaks that are strongly concentrated between 1.5 and 2.5 kHz. These features suggest a highly localized response to material irregularities in the defective area. At Point B, while the convoluted signal still contains distinct peaks in the same general frequency range, they appear broader and slightly less intense, indicating a spectral transition. By the time the signal reaches Point C, the convoluted spectrum becomes more dispersed and less pronounced, with peaks that are flatter and slightly shifted toward lower magnitudes. Though the frequency band remains consistent, the contrast and energy associated with specific frequencies diminish significantly.



(a)



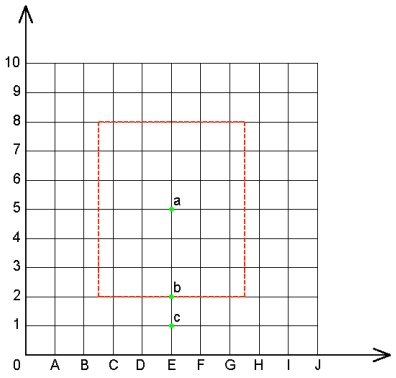
(b)

Figure 4. Frequency domain analysis of 200 mm width. (a) gridlines; (b) peak frequency of points.

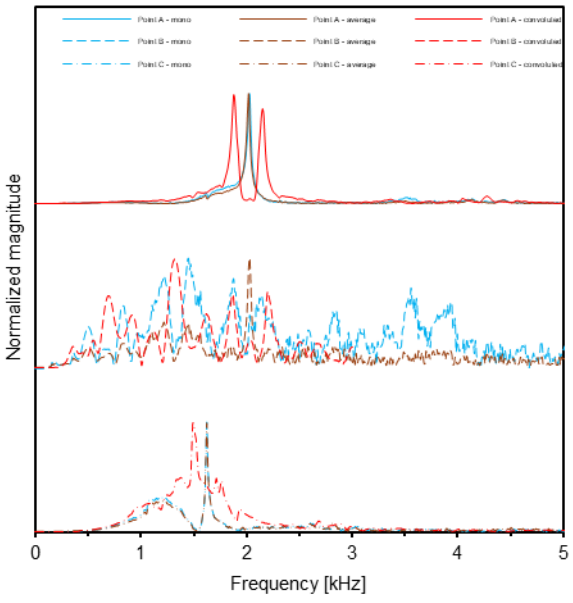
In the case of the average signal response, an opposite trend is observed. At Point A, the spectrum is marked by a single prominent and clearly distinguishable peaks, again concentrated in the 1.5–2.5 kHz region. Moving to Point B, subtle peak present at 1.5 kHz with the most prominent occurring at 2 kHz and smoothened frequency afterwards. The subtle peaks broaden and reduce in intensity, and at Point C, they appear smoother and more uniform. The average signal reflects a gradual flattening of spectral energy, with some redistribution toward a less defined frequency profile as one moves from the defect area to the sound region.

The mono signal response displays a subtler but consistent trend across the three points. At Point A, the mono spectrum is relatively flat, with minor irregularities and few notable peaks. At Point B, multiple subtle peaks are present at 1.5 kHz and 3.5 kHz with the most prominent occurring at 2 kHz. By Point C, the mono signal shows minimal spectral structure, with a consistent reduction in intensity and fewer identifiable peaks. Although no sharp frequency shift is evident, the general reduction in spectral complexity is apparent.

For 300 mm void width, a gradual shift in frequency content from Point A to Point C as illustrated in Figure 5 particularly in the convoluted signals. While Point A exhibits sharply defined peaks around 2 kHz, Point B displays broader and more distributed peaks, and Point C shows a more flattened, spread-out spectral profile. This shift suggests a transition from concentrated frequency energy in the defect region to a more diffuse and evenly distributed spectrum in the sound area. The dominant frequency components appear to migrate slightly lower and become less sharply localized as the spatial position moves away from the defect.



(a)



(b)

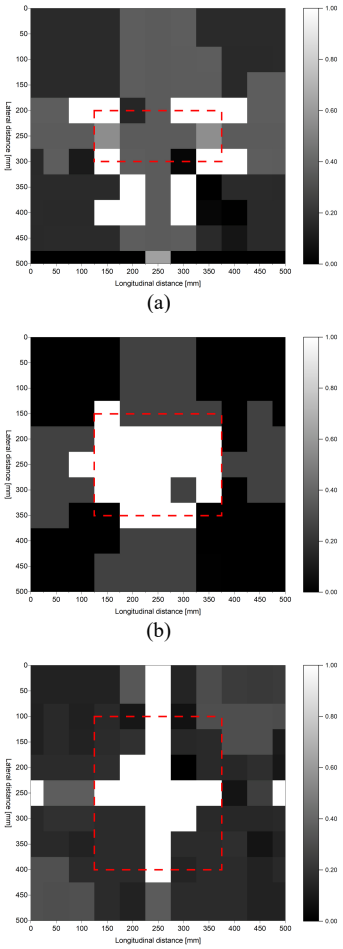
Figure 5. Frequency domain analysis of 300 mm width. (a) gridlines; (b) peak frequency of points.

At Point A, the spectrum shows very sharp, narrow peaks of high magnitude concentrated near 2 kHz, with minimal energy outside this range. Continuing to Point B, the spectrum becomes more dynamic, irregular, with scattered peaks and a more

evenly distributed energy pattern, suggesting increased spectral complexity compared to the defect region. At Point C, the mono signal displayed moderate peaks and broader frequency activity, particularly in the range of 1 to 2 kHz.

Average signal response displayed similar variation but with a lower magnitude at Point A. This high concentration of spectral energy is reduced at Point B, where the peaks become broader and span a wider frequency range—from approximately 1.5 to 2.5 kHz—resulting in a less focused frequency profile. By Point C, the spectral signature flattens further, showing broad, gentle peaks with lower amplitude and a more uniform energy distribution, particularly between 1 and 2 kHz.

In the case of convoluted signal response, Point A is marked by two narrow and high-intensity peaks centered around 2 kHz, highlighting the presence of distinct and localized frequency components. Point B retains the central frequency structure but exhibits broader, less intense peaks that suggest a dispersion of energy across nearby frequencies. Point C continues this trend, showing flatter and wider peaks with reduced magnitude, indicating a further spread and dilution of spectral energy.



(a) (b) (c)

Figure 6. Cavity localization (a) 100 mm, (b) 200 mm and (c) 300 mm.

The influence of signal responses on cavity localization is represented in Table 1 based on the localization of voids in Figure 6. The mono response signal exhibits varied performance across the different void widths. For the smallest void (100 mm), the detection rate was relatively low at 15%. This low detection rate implied that small voids are more difficult to detect using raw mono signals, possibly due to insufficient contrast or weak signal variations. Interestingly, as the void width increases to 200 mm, the detection rate improves substantially to 53%, indicating that larger voids produce more distinguishable signal features. However, at 300 mm, the rate decreases slightly to 40%, indicating a potential saturation point where further increase in void size does not necessarily translate to higher detectability.

Table 1. Detection rate (k) for signal responses across void widths

Signal responses	Width [mm]	k [%]
Mono	100	15
	200	53
	300	40
Average	100	5
	200	60
	300	37
Convoluted	100	10
	200	75
	300	37

Average signal response showed a different trend performing poorly for 100 mm with a detection rate of only 5%, the lowest among all categories. This could be due to the averaging process dampening subtle variations that are crucial for identifying small voids. However, its performance significantly improves by 200 mm with a detection rate of 60%, the highest among all methods for this width. The rate again drops to 37% for 300 mm voids. The sharp improvement from 100 mm to 200 mm suggests that averaging becomes effective when the voids are large enough to produce distinct combined signal patterns. Yet, the decline at 300 mm may indicate loss of fidelity, which reduces sensitivity to further increases in void size.

For the convoluted signal response, it demonstrated the most promising results overall. It achieves a 10% detection rate for 100 mm voids better than the average signal. The performance peaks impressively at 75% for 200 mm voids, suggesting that convolution significantly enhances detectability by highlighting pattern correlations or spatial features associated with voids. For 300 mm voids, the detection rate drops to 37%, mirroring the downward trend seen in the other methods. Nonetheless, the consistently high peak performance at the medium width implies that convolutional processing might be particularly effective at capturing structural signatures associated with voids of intermediate size.

The error analysis is based on the defect capability of the signal response and is shown in Figure 7. At 100 mm, the mono signal response exhibits the lowest error rate, indicating that raw or minimally processed data may retain critical high-frequency components and transient features needed for detecting small-scale voids. In contrast, both the average and convoluted signal responses report higher error rates. This performance gap may be attributed to the fact that averaging suppresses localized anomalies, potentially smoothing out the small perturbations caused by narrow voids. Similarly, convolution may amplify noise or introduce artifacts that degrade detection performance when spatial resolution is crucial (Liu et al., 2024).

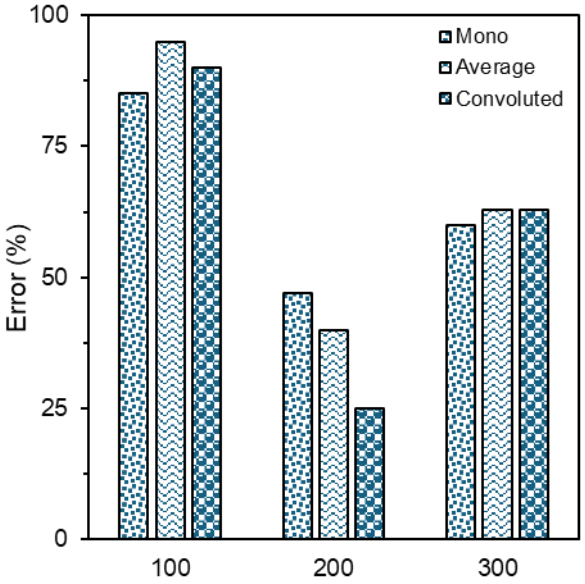


Figure 7. Void detection accuracy with varying width across several signal responses.

As the void width increases to 200 mm, a pivotal shift occurs. The convoluted signal response achieves the lowest error rate at this intermediate scale. This suggests that the convolution operation, often used to enhance signal features such as edges or energy signatures and benefits from a stronger acoustic reflection of larger voids (Gardner and Taylor, 2007; 2009; Mohler, 2014; Zhang, 2019). The signal-to-noise ratio improves, allowing convolution-based enhancement to emphasize structural discontinuities more effectively. The mono and average signal responses still perform adequately but with slightly elevated errors, potentially due to reduced sensitivity to mid-sized reflections when not specifically enhanced (Lander and Berbari, 1989; Ashley and Hassan, 2021; Wei, Xu and Zhang, 2024).

At 300 mm, the trend converges as all three signals show uniformly low error rates, indicating that voids of this size present a sufficiently large acoustic signature to be reliably detected regardless of the processing method. The performance plateau at this width indicates that once the void reaches a certain threshold in size, the dominant features captured in the signal are robust against variations in preprocessing strategy. The low intra-group variability observed across signal types at 300 mm highlights this consistency.

4 CONCLUSION

To improve the detectability of voids in structures using microphones, a study was conducted by introducing stereo microphones and signal processing techniques of average and convolution. FFT based analysis was used to detect cavities of varying widths in concrete structures. The following observations from the study are as follows:

a) The defect region amplifies and sharpens specific frequency components, the edge region softens and broadens them, and the sound area diffuses them into a more uniform and lower-energy profile especially for convoluted signal response.
b) Average signal response showed good frequency peak shift with increase in void size as compared to convoluted signal response which was prone to major fluctuations in the sound region especially for 300 widths.
c) Convoluted signal response clearly detected the voids with minor errors as compared to mono and average signal responses.
d) Voids of 100 mm were impossible to detect with detection rate less than 20%.

5 ACKNOWLEDGEMENTS

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (RS-2024-00339468).

6 REFERENCES

Akita, D., 2016. Multiplication and convolution of distributions for signal processing theory. *Digital Signal Processing*, 56, pp.1–14. https://doi.org/10.1016/j.dsp.2016.05.008.

Ashley, B.K. and Hassan, U., 2021. Time-domain signal averaging to improve microparticles detection and enumeration accuracy in a microfluidic impedance cytometer. *Biotechnology and Bioengineering*, 118(11), pp.4428–4440. https://doi.org/10.1002/bit.27910.

Chang, C.C., Yu, C.P. and Lin, Y., 2019. Distinction between crack echoes and rebar echoes based on Morlet Wavelet Transform of impact echo signals. *NDT and E International*, 108. https://doi.org/10.1016/j.ndteint.2019.102169.

Gardner, J.W. and Taylor, J.E., 2007. Novelconvolution Based Signal Processing Techniques for a Simplified Artificial Olfactory Mucosa. In: *TRANSDUCERS 2007 - 2007 International Solid-State Sensors, Actuators and Microsystems Conference*. IEEE. pp.2473–2476. https://doi.org/10.1109/SENSOR.2007.4300672.

Gardner, J.W. and Taylor, J.E., 2009. Novel convolution-based signal processing techniques for an artificial olfactory mucosa. *IEEE Sensors Journal*, 9(8), pp.929–935.

Gong, S., Feng, X., Zhang, G. and Ansari, F., 2023. Distributed detection of internal cavities in concrete-filled steel tube arch bridge elements. *Structural Health Monitoring*, 22(1), pp.657–671. https://doi.org/10.1177/14759217221088457.

Hassan, U. and Anwar, M.S., 2010. Reducing noise by repetition: introduction to signal averaging. *European Journal of Physics*, 31(3), pp.453–465. https://doi.org/10.1088/0143-0807/31/3/003.

Inaba, K., Tanigawa, H. and Naito, H., 2023. A study on evaluating supporting condition of railway track slab with impact acoustics and non-defective machine learning. *Construction and Building Materials*, 373, p.130905. https://doi.org/10.1016/j.conbuildmat.2023.130905.

Jiang, W., Xie, Y., Wu, J., Guo, J. and Long, G., 2021. Identifying bonding interface flaws in CRTS III type ballastless track structure using the impact-echo method. *Engineering Structures*, 227, p.111429. https://doi.org/10.1016/j.engstruct.2020.111429.

Kang, S., Yu, J.D., Hong, W.T. and Lee, J.S., 2021. Estimation of cavities beneath plate structures using a microphone: Laboratory model tests. *Sensors*, 21(9). https://doi.org/10.3390/s21092941.

Kuchipudi, S.T., Pudovikov, S., Wiggenhauser, H., Ghosh, D. and Rabe, U., 2023. Imaging of vertical surface-breaking cracks in concrete members using ultrasonic shear wave tomography. *Scientific Reports*, 13(1), p.21744. https://doi.org/10.1038/s41598-023-48699-w.

Lander, P. and Berbari, E.J., 1989. Optimizing signal averaging methods. In: *Images of the Twenty-First Century. Proceedings of the Annual International Engineering in Medicine and Biology Society,*. IEEE. pp.19–20.

Li, Bi., Cao, J., Xiao, J., Zhang, X. and Wang, H., 2014. Robotic Impact-echo Non-Destructive Evaluation based on FFT and SVM. In: *Proceeding of the 11th World Congress on Intelligent Control and Automation Shenyang, China*. pp.2854–2859.

Liu, J., Shi, C., Ge, L., Tie, R., Chen, X., Zhou, T., Gu, X. and Shen, Z., 2024. Enhanced Wind Field Spatial Downscaling Method Using UNET Architecture and Dual Cross-Attention Mechanism. *Remote Sensing*, 16(11). https://doi.org/10.3390/rs16111867.

Mohler, G.O., 2014. Learning convolution filters for inverse covariance estimation of neural network connectivity. Advances in Neural Information Processing Systems, 27.

Wei, J., Xu, C. and Zhang, S., 2024. Why Contour Averaging Works for SEM Metrology: Analysis and Validation. IEEE Transactions on Semiconductor Manufacturing, 37(4), pp.535–541. https://doi.org/10.1109/TSM.2024.3471635.

Xu, J. and Yu, X., 2021. Detection of Concrete Structural Defects Using Impact Echo Based on Deep Networks. Journal of Testing and Evaluation, 49(1), pp.109–120. https://doi.org/10.1520/JTE20190801.

Yao, F., Chen, G. and Abula, A., 2018. Research on signal processing of segment-grout defect in tunnel based on impact-echo method. Construction and Building Materials, 187, pp.280–289.

Yeh, P.-L., Liu, P.-L. and Hsu, Y.-Y., 2018. Parametric analysis of the impact-echo phase method in the differentiation of reinforcing bar and crack signals. Construction and Building Materials, 180, pp.375–381.

Yu, C.-P., Lin, Y. and Chang, C.-C., 2021. An effective crack-identification approach for impact echo signals using MWT spectrograms and scaled FFT spectra. Materials and Structures, 54, pp.1–21.

Yumnam, M., Ghosh, D. and Gupta, H., 2023. Empirical mode decomposition based techniques for imaging of shallow delamination in concrete using impact echo. Mechanical Systems and Signal Processing, 184, p.109668. https://doi.org/10.1016/j.ymssp.2022.109668.

Zhang, R., 2019. Making convolutional networks shift-invariant again. In: International conference on machine learning. PMLR. pp.7324–7334.