Journal of Asian Concrete Federation

journal homepage: https://jacf.sfulib3.publicknowledgeproject.org/index.php/jacf

Multi-Indicator Evaluation of Tile–Wall Integrity Based on Hammering Tests toward Automated Inspection

Yutong Li[*,a], Kei-ichi Imamoto[a], Chizuru Kiyohara[a], Masato Sato[a], Takaaki Ohkubo[b], Kazuya Sugiyama[c], Kohichi Umeda[c], Tatsuya Sugiyama[c]

[a] *Department of Architecture, Faculty of Engineering, Tokyo University of Science, 1, 6-3-1 Niijuku, Katsushika-ku, Tokyo 125-8585, Japan*
[b] *Department of Lifestyle Design, Yasuda Women's University, 6-13-1 Yasuhigashi, Asaminami-ku, Hiroshima, 731-0153, Japan*
[c] *SUGITEC CORPORATION Co., Ltd., 94 Nishida-cho, Nishiin, Ukyo-ku, Kyoto, 615-0055, Japan*

A B S T R A C T

This study investigates the bonding integrity of exterior wall tiles through hammering tests and multi-domain acoustic feature analysis. Three representative bonding conditions—intact tiles, tile–mortar debonding, and mortar–substrate debonding—were examined. A total of 22 quantitative indicators were extracted from the time-domain waveform, signal envelope, and frequency spectrum to characterize different debonding behaviors. The results show that energy- and amplitude-based indices, such as RMS, Energy, and Crest Factor, effectively distinguish intact and debonded states. Envelope-based parameters, including Envelope Area and Temporal Centroid, reveal the distribution and attenuation characteristics of vibration energy. Frequency-domain features, such as Spectral Centroid, Kurtosis, and Peak Power, capture the distribution, shape, and energy characteristics of the spectral response. The proposed multi-indicator framework provides a quantitative and comprehensive basis for evaluating tile adhesion conditions and contributes to the development of automated hammering-based inspection systems for building façades.

Keywords: Hammering Test, Time-Domain Feature, Envelope Analysis, Frequency-Domain Feature, Multi-Indicator Evaluation, Tile–Wall Bonding Integrity

1. Introduction

1.1 Background

The safety and durability of exterior wall finishing materials are directly related to urban public safety and the quality of living environments. However, due to factors such as material aging, construction defects, environmental influences, and increasing service life, the bonding condition between exterior tiles and their substrates may deteriorate over time [1]. Such deterioration not only affects the performance and aesthetics of buildings but can also pose serious safety hazards to pedestrians [2]. Therefore, rapid detection and quantitative evaluation of tile bonding integrity have attracted increasing attention in recent years. In response, many countries and regions have enacted relevant laws and technical standards that require periodic inspection of tile adhesion conditions to ensure building safety [3,4,5,6].

In current engineering practice, the hammering test has been widely used in Japan as a non-destructive, low-cost, and convenient method for the preliminary and secondary inspection of exterior wall tiles. The basic principle involves generating transient acoustic waves by striking the tile surface and judging the bonding condition (intact or debonded) based on the subjective perception of sound sharpness or dullness. However, conventional hammering inspection has inherent limitations in subjectivity, repeatability, and traceability. The results depend heavily on the experience of the inspector, and the criteria for distinguishing between bonding states are often inconsistent. Variations in impact force, striking

* Corresponding author: Yutong Li (a-yutong.li@rs.tus.ac.jp)

DOI http://dx.doi.org/10.18702/acf.2026.12.2.29

Received: 02-Nov-2025; Revised: 16-Feb-2026; Accepted: 23-Feb-2026; Published online: 21-May-2026

ISSN 2465-7964, eISSN 2465-7972. Copyright © Asian Concrete Federation, All rights reserved.

position, and ambient noise introduce further uncertainties [7,8]. In addition, the erection of scaffolding during large-scale inspections increases both cost and safety risks. With the growing demand for automated maintenance and the shortage of skilled labor in aging societies, robot- or drone-based automated hammering inspection systems have emerged as an important development direction [7].

To address these limitations, several researchers have developed automated and intelligent hammering inspection technologies in recent years. Luk et al. introduced several climbing robots and a novel automatic impact acoustic technique for detecting tiled walls, which combined mechanical mobility with intelligent sound analysis [4]. Sonoda et al. proposed a rotary hammering test method that significantly improved the accuracy and efficiency of conventional hammering inspections [9]. They evaluated its diagnostic capability through experiments on concrete specimens and acoustic analysis using the finite element method. Lin et al. designed a multi-species bionic climbing robot, which integrates an artificial intelligence–based acoustic recognition module achieving 74.62% accuracy in tile defect detection [10]. Zhao et al. developed a non-contact detection method for tile debonding on walls using a laser Doppler vibrometer combined with swept-band acoustic excitation, demonstrating high precision without physical impact [11]. Fujii et al. proposed a noise-robust classification approach based on an enhanced defect detector, showing high accuracy and environmental adaptability in concrete tunnel inspections [12]. Furthermore, Soeta et al. developed a prototype device equipped with an efficient impact mechanism and microphone, and investigated effective impact acoustic analysis methods, diagnostic indicators for debonding, and predictive approaches for estimating debonding depth [13].

To overcome the effects of surface roughness and ambient noise on detection, Tong et al. proposed a quantitative detection method based on time-domain acoustic features, together with a multilayer backpropagation artificial neural network classifier [14]. Yamamoto et al. also developed a small solenoid-valve impact device to address the influence of surrounding noise environments [15]. Their results showed that the rebound coefficient and rebound ratio could be used as effective indicators for diagnosing exterior wall tile debonding. Nagai et al. designed a drone system equipped with a squeegee impactor to avoid repeated strikes from spherical hammers [16]. They conducted acoustic detection of tile debonding and separated propeller noise from impact signals using wavelet analysis maps. Furthermore, Luk et al. developed a robust model-based inspection method that utilized wavelet-domain features and Hidden Markov Models (HMMs) to identify adhesive integrity [5]. They demonstrated that this method was highly resistant to variations in surface roughness and environmental noise.

In terms of evaluation indicators, Inoue et al. employed wavelet analysis and introduced an automated inspection robot for detecting tile debonding. They established a diagnostic method capable of quantitatively identifying two debonding modes: external and internal debonding [6,17]. Tanaka et al. visualized scalogram plots derived from wavelet analysis and attempted to estimate the depth and area of tile debonding by analyzing the maximum wavelet coefficients and reverberation time [18]. Emoto et al. utilized the Fast Fourier Transform (FFT) to extract frequency-domain features from tile sounding tests, compared the performance of multiple machine learning algorithms, and investigated the application potential of autoencoders in debonding detection [19]. Kitamura et al. determined the presence of debonding in artificial experimental specimens by analyzing variations in power spectra, and further introduced a vibrational model to estimate the direction and radius of debonding for different geometric configurations [20].

1.2 Purpose

Although previous studies have made notable progress in automation and intelligent inspection, there remains a lack of systematic analysis on the quantitative relationship between the physical characteristics of acoustic signals and the bonding condition of tiles. Moreover, most existing evaluation frameworks focus on single-dimensional features, and a universal indicator applicable to all inspection scenarios has not yet been established.

Therefore, this study proposes a multi-indicator acoustic feature analysis framework for evaluating the bonding integrity of exterior wall tiles. Field hammering tests were conducted on the exterior walls of a reinforced concrete apartment building to collect high-precision acoustic data representing three typical bonding conditions: intact (Intact), tile–mortar interface debonding (Debonded 1), and mortar–substrate interface debonding (Debonded 2). Through data preprocessing and signal analysis, a total of 22 quantitative indicators were extracted, including time-domain features (e.g., Maximum Absolute Value, Root Mean Square, Energy, and Crest Factor) representing impact intensity; envelope-based features characterizing energy decay and persistence; and frequency-domain features (e.g., Spectral Centroid, Roll-off, Skewness, and Kurtosis) describing the influence of bonding conditions on frequency distribution.

Comparative analysis revealed clear differences between intact and debonded samples in terms of energy distribution, decay characteristics, frequency concentration, and spectral symmetry. Furthermore, several indicators such as Bandwidth and Spectral Flatness demonstrated the potential to distinguish between different debonding modes. This study establishes a systematic multi-indicator acoustic analysis framework that provides a reliable foundation for the future development of machine-learning-based automated tile inspection systems.

2. Hammering Test

The hammering test was conducted on a ten-story reinforced concrete apartment building located in Tokyo, which was completed in 2006. The exterior walls are finished with ceramic tiles. In this study, the hammering test was performed on the south façade of the fifth floor to obtain

acoustic data representing different bonding conditions.

Prior to the measurements, experienced inspectors conducted a conventional hammering test to classify the wall surface into three typical conditions: (1) Intact tile (Intact) – tiles with normal adhesion; (2) Tile–mortar interface debonding (Debonded 1) – debonding occurring at the interface between the ceramic tile and the finishing mortar; (3) Mortar–substrate debonding (Debonded 2) – debonding at the interface between the finishing mortar and the concrete substrate. An overview and corresponding photographs of the hammering test are shown in Fig. 1 and Photo 1, respectively.

From each of the three regions, 10 intact tiles, 8 tiles with Debonded 1, and 8 tiles with Debonded 2 were randomly selected for hammering tests. Each tile was struck three times using a commercially available hammering rod. The hammering was performed manually without strict quantitative control of the impact force; however, the operator attempted to maintain approximately consistent drop distances between impacts according to typical inspection practice. The reflected sounds were recorded using an AT4022 instrument microphone and a ZOOM F3 field recorder, with a sampling frequency of 96 kHz and stored in 32-bit floating-point WAV format.

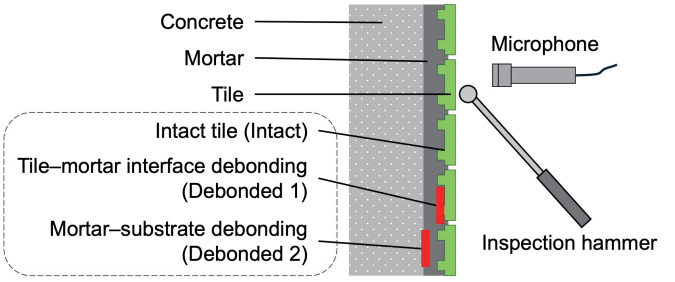


Fig. 1. Overview of the hammering test



(a) Microphone and inspection hammer (b) Target tiles

Photo 1. Experimental photos

3. Data Preprocessing and Feature Extraction

For each tile, the three hammering sounds were separated and individually analyzed. The average value of the three strikes was used to represent the acoustic characteristics of each tile in the subsequent feature extraction process.

The original 32-bit floating-point WAV files were converted into CSV format for detailed analysis. As shown in Fig. 2, signal segmentation was performed by detecting abrupt changes in the waveform using a standard deviation (STD) threshold of 10 to identify the onset point of each impact, which was empirically determined through pilot tests as it provided stable and reliable onset detection for the hammering signals. For each detected impact, a 0.1 s segment was extracted, starting 0.02 s before the onset point, to capture the full hammering response, where the segment length was selected to ensure that the complete impact response was included while avoiding unnecessary redundancy in the signal data. After the end of the first impact was identified, the search was continued to detect the next onset point, and the procedure was repeated until all three impacts were extracted. Subsequently, acoustic feature extraction was conducted in three analytical domains: (1) time-domain waveform; (2) envelope curve; (3) frequency-domain spectrum. A total of 22 quantitative features were computed across these domains to characterize and distinguish the bonding conditions of tiles.

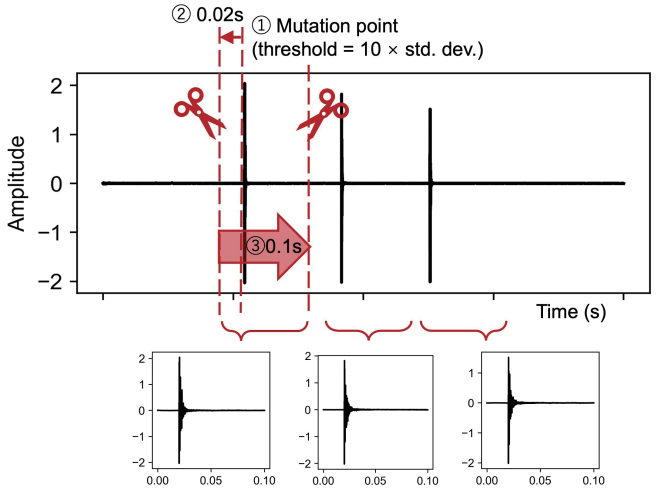


Fig. 2. Extraction of the sound waveform segment

3.1 Time-Domain Features and Their Calculation Methods

To quantitatively analyze the hammering sound characteristics under different debonding conditions, four typical time-domain features were extracted from the raw waveform data: Maximum Absolute Value, Root Mean Square (RMS), Energy, and Crest Factor. The definitions and calculation methods of these features are as follows.

(1) Maximum Absolute Value

The Maximum Absolute Value represents the instantaneous peak amplitude of the hammering response, which reflects the impact intensity of the signal. It is calculated as:

$$A_{max} = \max(|x[n]|) \quad (1)$$

where x[n] is the discrete waveform signal at the n-th

sampling point. The calculation was performed for a 0.1 s segment (9600 sampling points).

(2) Root Mean Square (RMS)

The RMS value represents the effective amplitude of the signal and reflects its overall magnitude. It is defined as:

$$x_{RMS} = \sqrt{\frac{1}{N}\sum_{n=0}^{N-1} x^2[n]} \quad (2)$$

where x[n] is the signal amplitude at the n-th sample and N is the number of sampling points in the analysis window (9600 in this study).

(3) Energy

The Energy of the signal corresponds to the cumulative sum of the squared amplitudes, indicating the signal's overall strength and duration. It is expressed as:

$$E = \sum_{n=0}^{N-1} x^2[n] \quad (3)$$

where x[n] is the discrete waveform amplitude and N is the total number of samples within the analysis window, which is consistent with that used in the RMS calculation.

(4) Crest Factor

The Crest Factor describes the sharpness or impulsiveness of the waveform and is defined as the ratio of the peak amplitude to the RMS value:

$$CF = \frac{A_{max}}{x_{RMS}} \quad (4)$$

where A_{max} is the maximum absolute amplitude and x_{RMS} is the RMS value of the signal. A higher Crest Factor indicates a stronger transient component and greater impulsive energy within the signal.

3.2 Envelope-Based Features and Their Calculation Methods

To further analyze the amplitude attenuation characteristics and temporal distribution of the hammering sounds, two types of signal envelopes were derived from the raw waveform: the peak-interpolated envelope and the Hilbert envelope. Six envelope-based features were then calculated to describe the decay behavior and energy characteristics of the signals.

(1)Peak-Interpolated Envelope and Its Decay Coefficient

For each hammering signal, a 0.1 s segment (9600 sampling points) starting from the peak amplitude was extracted and amplitude-normalized. Local maxima and minima were detected to represent the positive and negative peaks of the waveform. Upper and lower envelopes were then constructed by linearly interpolating between successive peaks and troughs. The time axis was linearly mapped to a range of 0–100 ms, and each envelope was fitted with an exponential decay function of the following form:

$$y = ae^{-bx} + c \quad (5)$$

where y is the fitted envelope amplitude, x is time in milliseconds, a is the initial amplitude, b is the amplitude decay coefficient, and c is the steady-state offset term. The upper and lower envelopes were fitted separately to obtain the corresponding decay coefficients b_{upper} and b_{lower}. The Decay Coefficient b reflects the rate at which the signal amplitude decreases over time—a larger value indicates faster attenuation and a shorter resonance duration. The mean value of b_{upper} and b_{lower} was used as the representative feature for each impact.

(2) Hilbert Envelope and Its Statistical Feature Indices

To describe the overall energy distribution and temporal evolution of the hammering signal, the Hilbert transform was employed to compute a smooth and continuous analytic envelope. The Hilbert envelope reflects the instantaneous amplitude of the signal and effectively characterizes the gradual decay of vibration energy over time. The analytic signal z(t) and its envelope e(t) are defined as follows:

$$z(t) = x(t) + j\,\mathrm{Hilbert}(x(t)) \quad (6)$$

$$e(t) = |z(t)| = |\mathrm{Hilbert}(x(t))| \quad (7)$$

where x(t) is the original waveform of the hammering sound, z(t) is the complex analytic signal obtained through the Hilbert transform, j is the imaginary unit, and e(t) represents the instantaneous amplitude of the signal at time t.

Based on the Hilbert envelope e(t), five statistical feature indices were extracted to quantify the signal's temporal and energetic characteristics. These features include the Envelope Area, Early Energy Ratio, Fluctuation Coefficient, Temporal Centroid, and Decay Time.

1) Envelope Area

The Envelope Area (A) represents the accumulated energy of the signal over the entire duration of the envelope. It is calculated as:

$$A = \int_0^T e(t)\,dt \quad (8)$$

where e(t) is the Hilbert envelope amplitude and T=100 ms is the duration of analysis. A larger area indicates a higher overall energy content in the hammering response.

2) Early Energy Ratio

The Early Energy Ratio (R_{early}) quantifies the concentration of energy in the early stage of the response. In this study, it is defined as the ratio of the energy within the first 3 ms, which is adopted here as a representative early-time window based on empirical observation of the signal response, to the total envelope area:

$$R_{early} = \frac{\int_0^{3\,ms} e(t)\,dt}{A} \quad (9)$$

where A is the total envelope area. A higher ratio implies that the majority of the energy is released at the initial impact, whereas a smaller value indicates a slower decay or sustained vibration.

3) Fluctuation Coefficient

The Fluctuation Coefficient (F) evaluates the stability of the envelope shape and is defined as the ratio of the root mean square to the mean value of the envelope:

$$F = \frac{\sqrt{\frac{1}{N}\sum_{n=0}^{N-1} e^2[n]}}{\frac{1}{N}\sum_{n=0}^{N-1} e[n]} \quad (10)$$

where e[n] is the discrete envelope amplitude and N is the number of envelope samples (9600 points in this study). A higher F value indicates larger fluctuations in the envelope, whereas a lower value corresponds to a smoother decay behavior.

4) Temporal Centroid

The temporal centroid (C_t) represents the center of gravity of the signal energy distribution over time and is defined as:

$$C_t = \frac{\int_0^T t\, e(t)\, dt}{\int_0^T e(t)\, dt} \quad (11)$$

where t is time and e(t) is the Hilbert envelope amplitude. In this study, the analysis range was T=100 ms. A smaller C_t indicates that most of the energy is concentrated in the early portion of the signal, whereas a larger C_t suggests a more prolonged response.

5) Decay Time

The Decay Time (T_{decay}) characterizes the rate at which the signal energy diminishes over time and reflects the duration of resonance within the structure. The Hilbert envelope e(t) is used to determine the peak value e_{max} and its corresponding index n_{max}. The Decay Time is defined as the duration required for the envelope to drop to 10% of its peak value, which was selected to capture the main energy decay stage of the signal while reducing the influence of low-amplitude noise in the tail of the envelope:

$$T_{decay} = \frac{n_{(10\%)} - n_{max}}{F_s} \quad (12)$$

where F_s=96,000 Hz is the sampling frequency, n_{max} is the index of the envelope's peak amplitude, and $n_{(10\%)}$ is the first index at which the envelope falls below $0.1e_{max}$.

3.3 Frequency-Domain Features and Their Calculation Methods

To reveal the differences in frequency distribution of hammering sounds under different debonding conditions, Fast Fourier Transform (FFT) was applied to the raw waveform signals. A total of twelve representative frequency-domain features were extracted to quantitatively describe the spectral characteristics of the signals. The selection of signal processing parameters and feature definitions was partly informed by commonly used approaches in audio signal analysis [21]. For each impact, the time-domain data corresponding to the first 2048 sampling points were used for analysis. Before applying FFT, the DC component was removed to eliminate measurement bias. The frequency-domain spectrum X(f) was then obtained by performing FFT on the time-domain signal x(t):

$$X(f) = FFT\{x(t)\} \quad (13)$$

The power spectrum P(f) was calculated as the squared magnitude of the normalized complex spectrum:

$$P(f) = \left|\frac{X(f)}{N}\right|^2 \quad (14)$$

where N is the number of sampling points. Only the positive-frequency range (0–48 kHz) was retained for subsequent analysis, since the FFT spectrum is symmetric about zero frequency.

In this study, twelve frequency-domain features were calculated and grouped into four categories according to their spectral characteristics:

Group 1: Features describing the concentration of spectral energy — Spectral Centroid, Spectral Bandwidth, and Spectral Roll-off;

Group 2: Features characterizing spectral complexity and smoothness — Spectral Entropy, Spectral Flatness, and High-Frequency Energy Ratio;

Group 3: Features representing the asymmetry and sharpness of the spectrum — Spectral Skewness and Spectral Kurtosis;

Group 4: Features quantifying the overall intensity and spectral trend — Peak Frequency, Peak Power, Spectral Slope, and Spectral RMS Power.

(1) Spectral Centroid

The Spectral Centroid (f_c) represents the center of gravity of the power spectral energy and indicates where most of the signal energy is concentrated within the frequency range. It is defined as:

$$f_c = \frac{\sum f_i P_i}{\sum P_i} \quad (15)$$

where f_i is the frequency at the i-th spectral bin, and P_i is the corresponding spectral power. A higher centroid value indicates a greater proportion of high-frequency components in the signal.

(2) Spectral Bandwidth

The Spectral Bandwidth (BW) measures the dispersion of the spectral energy around the centroid, reflecting the effective width of the main frequency band. It is calculated as:

$$BW = \sqrt{\frac{\sum (f_i - f_c)^2 P_i}{\sum P_i}} \quad (16)$$

where f_i is the frequency, f_c is the spectral centroid, and P_i is the spectral power. A larger bandwidth indicates that the energy is more widely distributed across the frequency range.

(3) Spectral Roll-off

The Spectral Roll-off Frequency (f_{roll}) is defined as the frequency below which a specified percentage of the total spectral energy is contained. In this study, the threshold was set to 85% of the total energy, which is a commonly used value in audio signal analysis to characterize the distribution of spectral energy (e.g.: [22]):

$$\sum_{f \le f_{roll}} P(f) = 0.85 \sum P(f) \quad (17)$$

where P(f) is the power spectrum. A higher roll-off frequency indicates that more energy is distributed in the higher-frequency region.

(4) Spectral Entropy

The Spectral Entropy (H) quantifies the irregularity or randomness of the spectral energy distribution. It is based on the normalized power spectrum $p_i=P_i/\sum P_i$ and is calculated as:

$$H = -p_i \log \sum (p_i) \tag{18}$$

where P_i is the spectral power at frequency f_i. A higher entropy value indicates that the spectral energy is more evenly distributed (i.e., more disordered), while a lower value suggests that the energy is concentrated within a few frequency bands.

(5) Spectral Flatness

The spectral flatness (F_{flat}) evaluates how flat or peaked a spectrum is and indicates how similar the spectral shape is to that of white noise. It is defined as the ratio of the geometric mean to the arithmetic mean of the power spectrum:

$$F_{flat} = \frac{\text{Geometric mean [P(f)]}}{\text{Arithmetic mean [P(f)]}} \tag{19}$$

where P(f) is the power spectrum. A higher flatness value indicates a flatter spectrum, meaning the signal contains more uniformly distributed frequency components (noise-like), while a lower value corresponds to a more tonal or peaked spectrum.

(6) High-Frequency Energy Ratio

The High-Frequency Energy Ratio ($R_{>3000}$) represents the proportion of energy contained in the higher frequency range (>3000 Hz) relative to the total energy, reflecting the perceived sharpness or brightness of the sound. The threshold of 3000 Hz was selected based on observations from the current dataset, in which this frequency range showed a clear distinction between debonded and intact tiles. It is given by:

$$R_{>3000} = \frac{\sum_{f>3000} P(f)}{\sum P(f)} \tag{20}$$

where P(f) is the spectral power at frequency f. A larger $R_{>3000}$ indicates stronger high-frequency content, corresponding to a sharper acoustic response.

(7) Spectral Skewness

The Spectral Skewness (S_k) measures the asymmetry of the spectral distribution relative to its mean value. It is defined as:

$$S_k = \frac{E[(P_i - \bar{P})^3]}{(\text{std}(P))^3} \tag{21}$$

where P_i is the spectral power at frequency f_i, $\bar{P}$ is the mean spectral power, and std(P) is the standard deviation of the power spectrum. A positive skewness (right-skewed) indicates that more energy is distributed in the high-frequency range, while a negative skewness (left-skewed) suggests dominance of low-frequency components. A value near zero implies a symmetrical energy distribution.

(8) Spectral Kurtosis

The Spectral Kurtosis (K) characterizes the sharpness or peakedness of the spectral distribution, indicating whether the spectrum contains a dominant narrowband peak or a broad, flat energy distribution. It is mathematically defined as the ratio of the fourth central moment to the squared variance:

$$K = \frac{E[(P_i - \bar{P})^4]}{(\text{std}(P))^4} \tag{22}$$

where P_i is the power spectrum value at frequency bin i, $\bar{P}$ is the mean power spectrum value, E[.] denotes the expectation operator, and std(P) represents the standard deviation of the power spectrum. A higher kurtosis value indicates the presence of a sharp spectral peak, meaning that the signal energy is concentrated in a narrow frequency band, while a lower kurtosis corresponds to a flatter spectrum with less pronounced peaks.

(9) Peak Frequency

The Peak Frequency (f_{peak}) corresponds to the frequency at which the power spectrum reaches its maximum value. It represents the dominant frequency component of the signal and is defined as:

$$f_{peak} = \arg\max\left(P(f)\right) \tag{23}$$

where P(f) is the power spectrum at frequency f. A higher Peak Frequency typically indicates stronger high-frequency content in the hammering response.

(10) Peak Power

The Peak Power (P_{peak}) is defined as the maximum amplitude of the power spectrum, representing the strongest frequency component within the signal:

$$P_{peak} = \max\left(P(f)\right) \tag{24}$$

where P(f) is the power spectrum at frequency f. This value reflects the intensity of the dominant resonance frequency in the spectral domain.

(11) Spectral Slope

The Spectral Slope (S) quantifies the rate at which the spectral amplitude increases or decreases with frequency, indicating the overall trend of energy distribution between low and high frequencies. It is expressed as:

$$S = \frac{\sum(f_i - \bar{f})(P_i - \bar{P})}{\sum(f_i - \bar{f})^2} \tag{25}$$

where f_i is the frequency, P_i is the corresponding spectral power, $\bar{f}$ is the mean frequency, and $\bar{P}$ is the mean power. A negative slope suggests that energy is concentrated in lower frequencies, while a positive slope indicates dominance of higher-frequency components.

(12) Spectral RMS Power

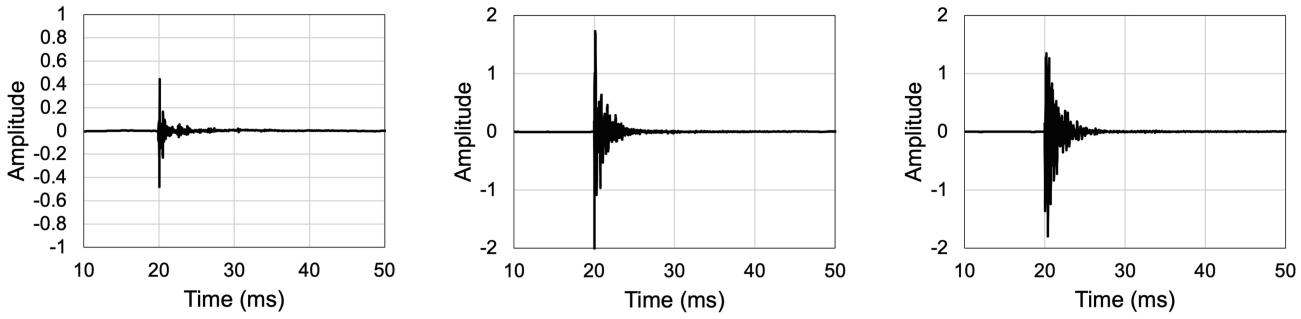
The Spectral Root Mean Square (RMS) power (P_{RMS}) represents the overall energy level of the power spectrum and is defined as the root mean square of the spectral power values:

$$P_{RMS} = \sqrt{\frac{1}{N}\sum P_i^2} \tag{26}$$

where N is the total number of frequency bins. A larger P_{RMS} value corresponds to higher overall spectral energy in the hammering signal.

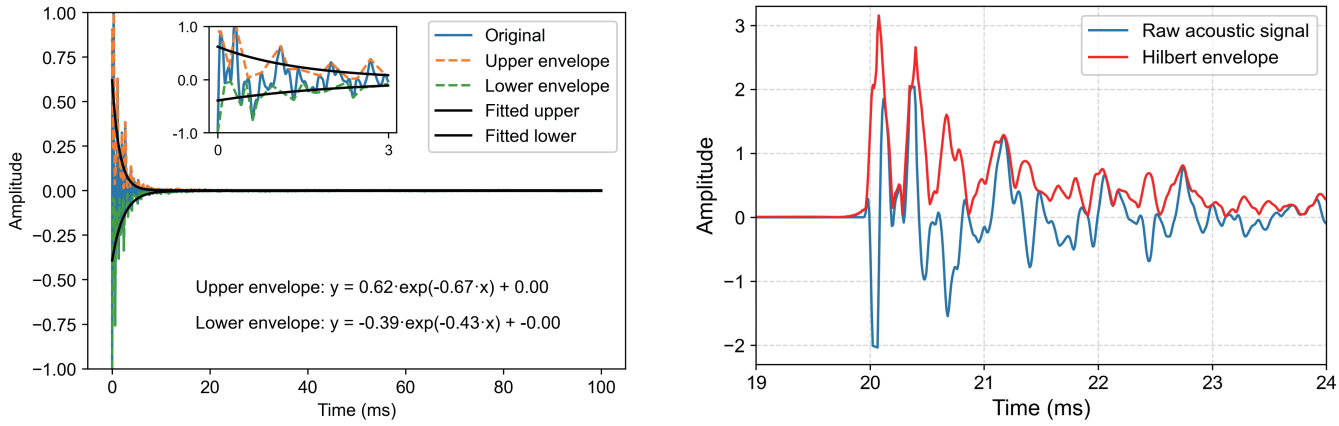
4. Results and Discussion

4.1Representative Result Example



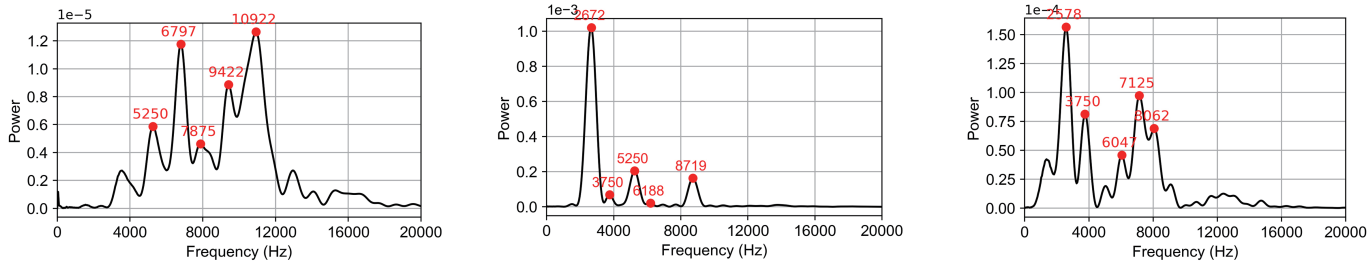
(a) Intact (b) Debonded 1 (c) Debonded 2

Fig. 3. Representative results of time-domain waveforms of hammering sounds



(a) Peak-Interpolated Envelope (b) Hilbert Envelope

Fig. 4. Representative results of envelope-based analysis of hammering sounds



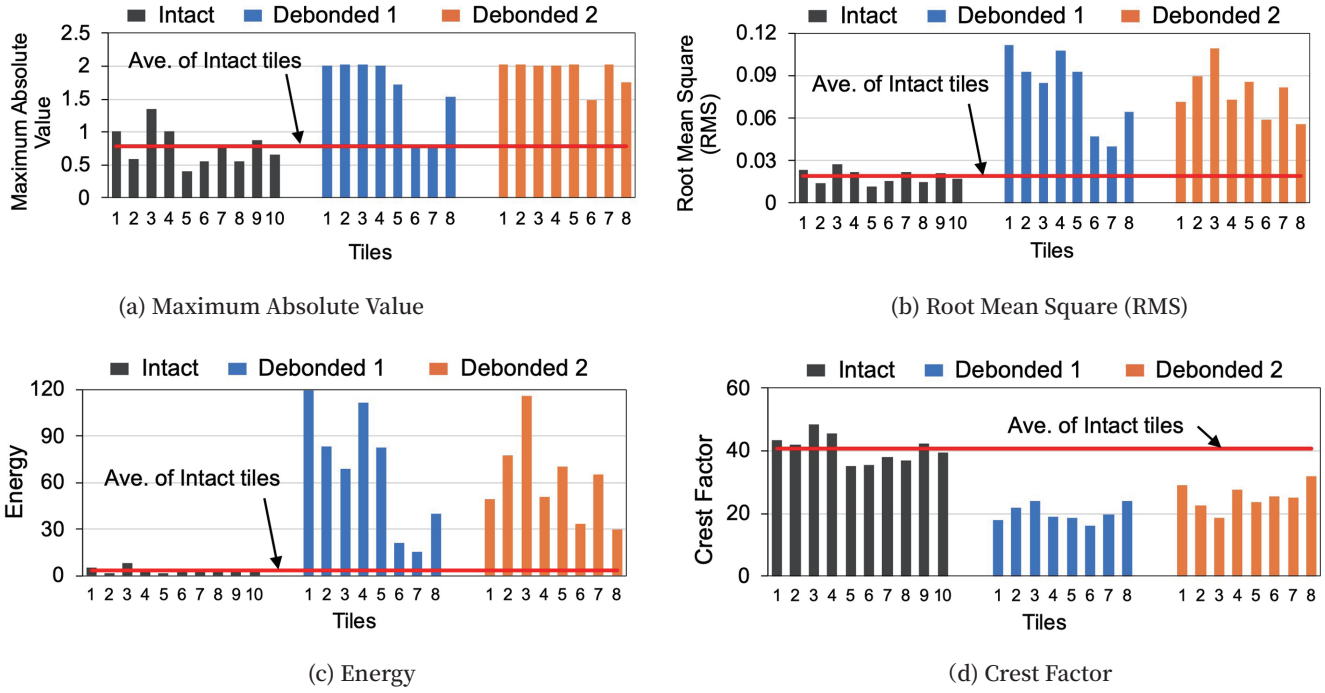
(a) Intact (b) Debonded 1 (c) Debonded 2

Fig. 5. Representative results of frequency-domain spectra of hammering sounds

To analyze the differences in acoustic responses under various bonding conditions, this study performed quantitative feature extraction and comparative analysis on three categories of samples: intact tiles (Intact), tile–mortar interface debonding (Debonded 1), and mortar–substrate interface debonding (Debonded 2). Fig. 3–5 present representative examples of the time-domain waveform, envelope-based analysis, and frequency-domain power spectrum, respectively, while Fig. 6–8 summarize the calculated feature indices for each category.

As shown in Fig. 3, the time-domain waveform of the Intact sample exhibits smaller amplitudes and rapid attenuation. In contrast, the debonded samples show significantly higher amplitudes and longer vibration durations, indicating stronger resonance behavior. Among them, Debonded 2 displays the slowest energy decay, suggesting more pronounced interface reflections and multiple resonance phenomena.

Fig. 4 (a) shows the upper and lower envelopes obtained by peak-interpolation along with their exponential decay fitting curves, whereas Fig. 4 (b) presents the analytic envelope derived from the Hilbert transform. The peak-interpolated envelope effectively captured the local peak variations and amplitude-decay trends of the waveform. It is more sensitive to instantaneous impact characteristics and is suitable for describing the rate of amplitude decay and resonance. However, it is also more susceptible to local noise and sampling density, which required precise identification of the fitting starting point. In contrast, the Hilbert envelope—a smooth



(a) Maximum Absolute Value

(b) Root Mean Square (RMS)

(c) Energy

(d) Crest Factor

Fig. 6. Results of time-domain acoustic features

and continuous curve obtained from the magnitude of the analytic signal—provides a stable representation of the overall energy distribution and temporal decay of the signal and is less affected by local fluctuations. Thus, this envelope is thus more appropriate for describing the overall energy-attenuation process and the temporal concentration of vibration energy. In summary, the peak-interpolation envelope emphasizes the "amplitude-decay dynamics," whereas the Hilbert envelope highlights the "energy-distribution smoothness." The combined analysis of both envelopes enables a comprehensive characterization of the vibration behavior of the tile-wall system.

Fig. 5 presents the representative frequency-domain power spectra for the three bonding conditions. For Intact tiles, the dominant peak frequencies are concentrated within approximately 5–11 kHz, indicating that their acoustic response is dominated by high-frequency components with a relatively balanced multi-peak distribution. In contrast, the Debonded 1 samples exhibit lower dominant frequencies (approximately 2–8 kHz) accompanied by a clear reduction in high-frequency energy. Furthermore, the Debonded 2 samples show two strong peak regions at 2–4 kHz and 6–8 kHz, implying the formation of local resonant cavities at the debonded interfaces, which leads to more complex energy-transfer paths and stronger reflection superposition effects.

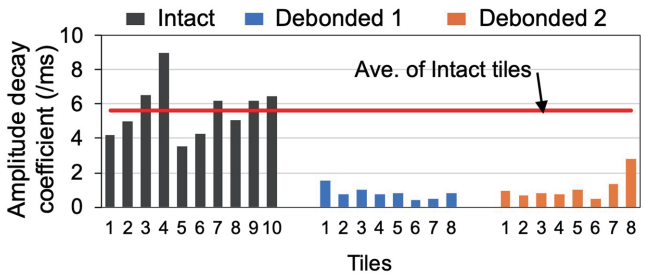
4.2 Time-Domain Feature Results

As shown in Fig. 6, the time-domain features exhibit clear distinctions among the three bonding conditions. The Maximum Absolute Value (Fig. 6 (a)) differs significantly between the intact and debonded regions. Overall, the debonded tiles show peak amplitudes that are more than twice those of the intact tiles on average; however, two tiles in the Debonded 2 group display values comparable to the intact ones. This may be attributed to variations in manual hammering force, indicating that the impact strength can influence the maximum amplitude. In this study, hammering was performed manually without strict quantitative control of the impact force; however, the operator attempted to maintain approximately consistent drop distances according to typical inspection practice to reduce variability between impacts. Therefore, under conditions where the impact force is kept consistent, this indicator can more accurately represent the acoustic response characteristics corresponding to different bonding states. Despite the manual excitation, the maximum amplitude still shows clear differences between bonding conditions. With further improvement, such as using mechanical excitation to better control the impact force, this indicator may provide even more reliable discrimination of bonding states.

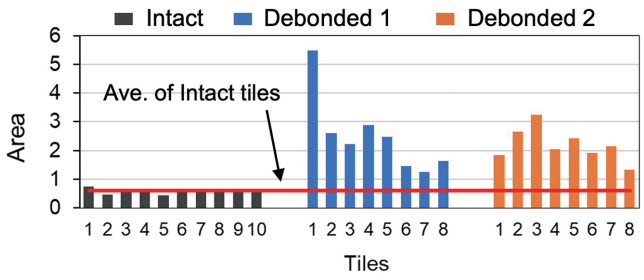
In addition, the Root Mean Square (RMS), Energy, and Crest Factor also show clear distinctions among bonding conditions. The RMS (Fig. 6 (b)) and Energy (Fig. 6 (d)) of the debonded regions are significantly higher than those of the intact ones, indicating stronger overall vibration energy. Conversely, the Crest Factor (Fig. 6 (c)) of debonded tiles is relatively lower, suggesting that their signals contain weaker instantaneous impact components and exhibit slower energy decay. In contrast, the intact tiles exhibit higher Crest Factors, implying sharper responses dominated by instantaneous impact components.

4.3 Envelope-Based Feature Results

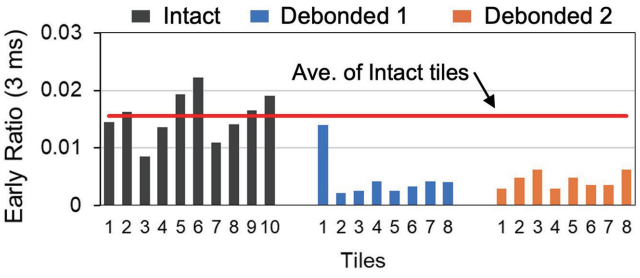
As shown in Fig. 7, the envelope-based features reveal



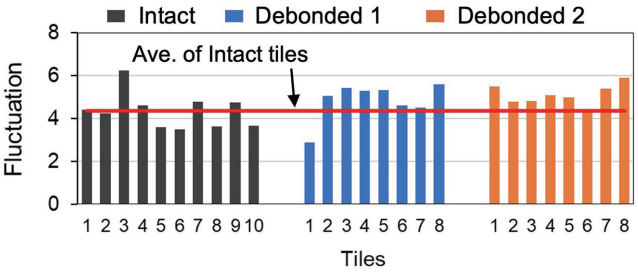
(a) Amplitude Decay Coefficient of the peak-interpolated envelopes



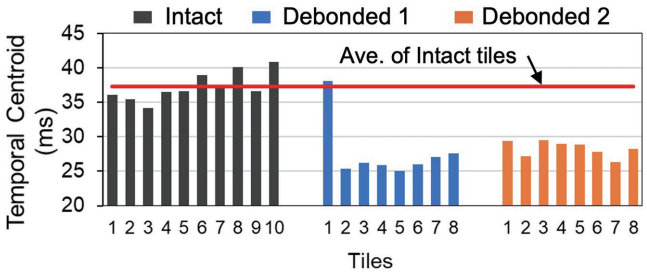
(b) Envelope Area (Hilbert envelope)



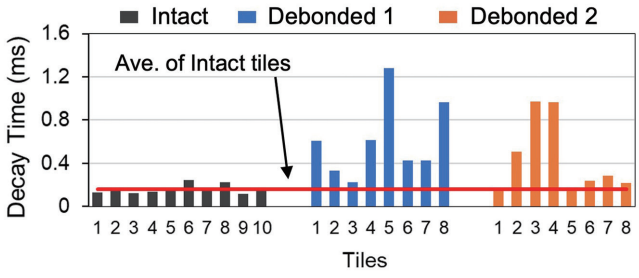
(c) Early Energy Ratio (Hilbert envelope, 3ms)



(d) Fluctuation (Hilbert envelope)



(e) Temporal Centroid (Hilbert envelope)

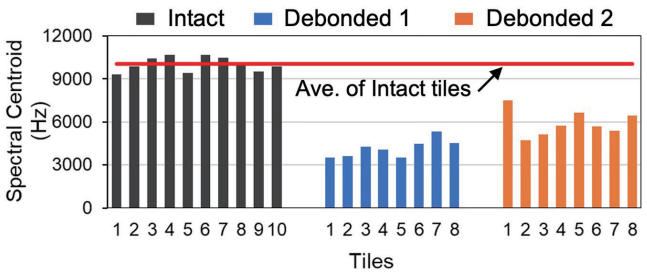


(f) Decay Time (Hilbert envelope, 10%)

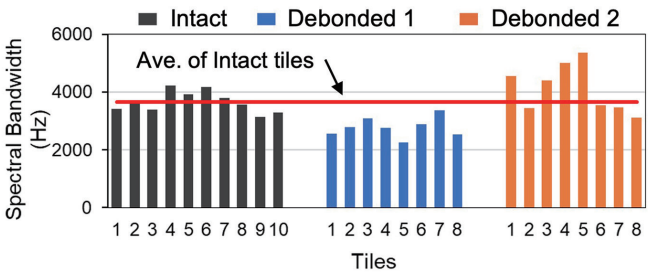
Fig. 7. Results of envelope-based acoustic features

distinctive differences among the three bonding conditions. The amplitude Decay Coefficient of the peak-interpolated envelope (Fig. 7 (a)) is highest for the Intact samples (with an average value of 5.63/ms), indicating faster energy attenuation and greater structural damping. In contrast, the debonded regions exhibit smaller Decay Coefficients (all values being below 2.83/ms), suggesting slower energy decay and stronger resonance persistence.

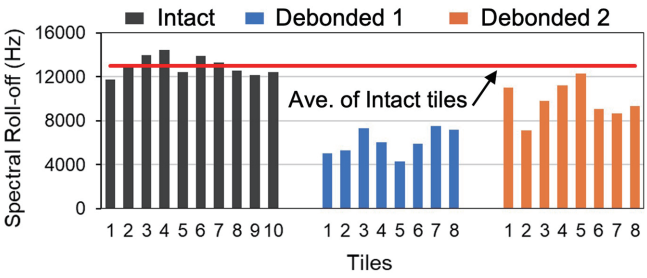
The Hilbert Envelope Area (Fig. 7 (b)) and Early Energy Ratio (Fig. 7 (c)) describe the temporal concentration of acoustic energy. Results show that the debonded regions possess larger Envelope Areas (the average values of Debonded 1 and Debonded 2 are about 3.9 times that of the Intact samples) and lower Early Energy Ratios (the average values of Debonded 1 and Debonded 2 are about 0.29 times that of the Intact samples), indicating a more gradual and prolonged energy release process. This tendency reflects that the reduced stiffness and multiple reflections at debonded interfaces cause the acoustic energy to dissipate over a longer time rather than being concentrated in the initial impact period. However, for the Fluctuation Coefficient (Fig. 7 (d)), no clear distinction is observed among the bonding conditions. This may be because the coefficient primarily reflects the overall smoothness of the envelope rather than local energy variations or decay characteristics. Hence, this indicator is less sensitive to debonding-related changes and cannot effectively discriminate between intact and debonded tiles. Regarding the Temporal Centroid (Fig. 7 (e)), most debonded samples exhibit smaller values than the average of the intact ones, suggesting that their energy is concentrated in the earlier stage of the signal. However, one Debonded 1 sample remains slightly above the intact average line. This might be attributed to the calculation range of the Temporal Centroid (currently 0–0.1 s). Adjusting the analysis window or adopting a weighted temporal integration could potentially enhance its discriminative ability, making it a useful supplementary indicator for debonding diagnosis. As for the Decay Time (Fig. 7 (f)), no consistent trend is observed in the current results. Nevertheless, this does



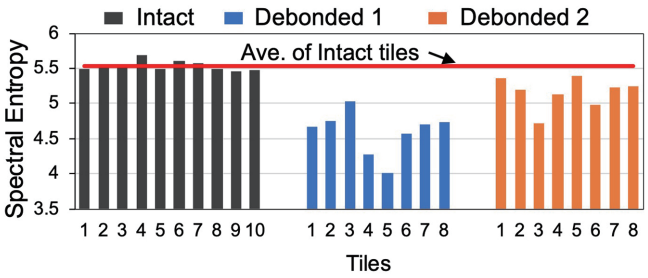
(a) Spectral Centroid



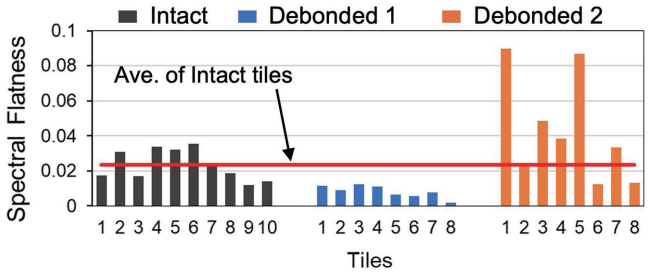
(b) Spectral Bandwidth



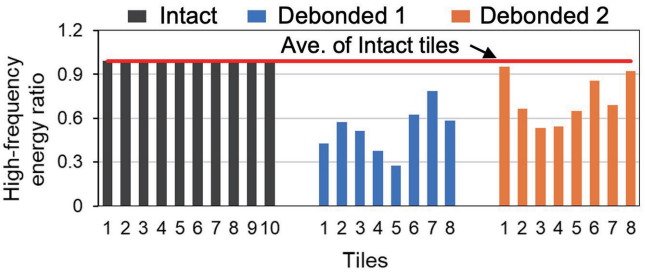
(c) Spectral Roll-off (85%)



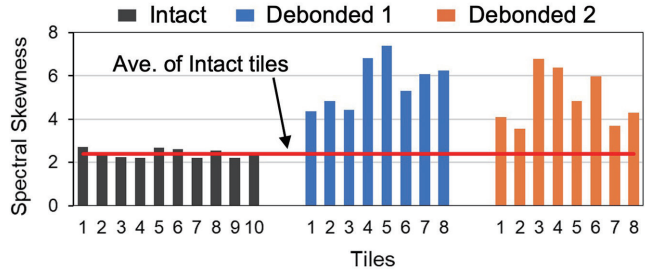
(d) Spectral Entropy



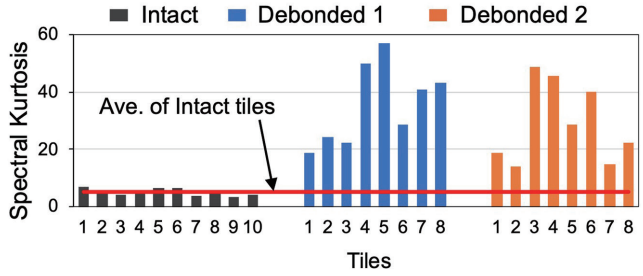
(e) Spectral Flatness



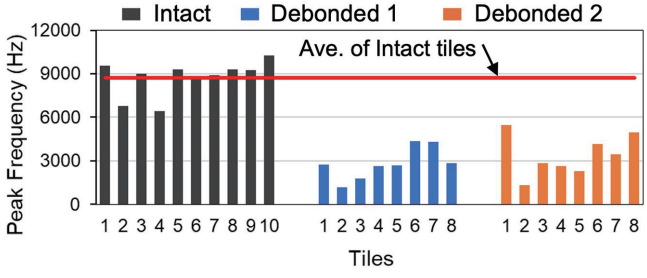
(f) High-Frequency Energy Ratio (>3000 Hz)



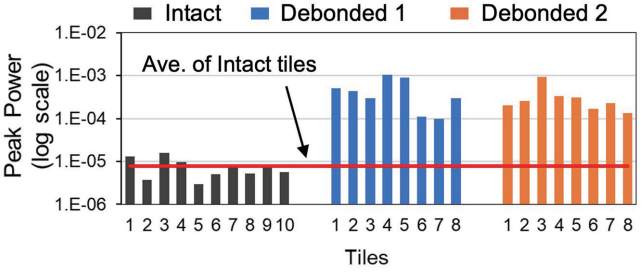
(g) Spectral Skewness



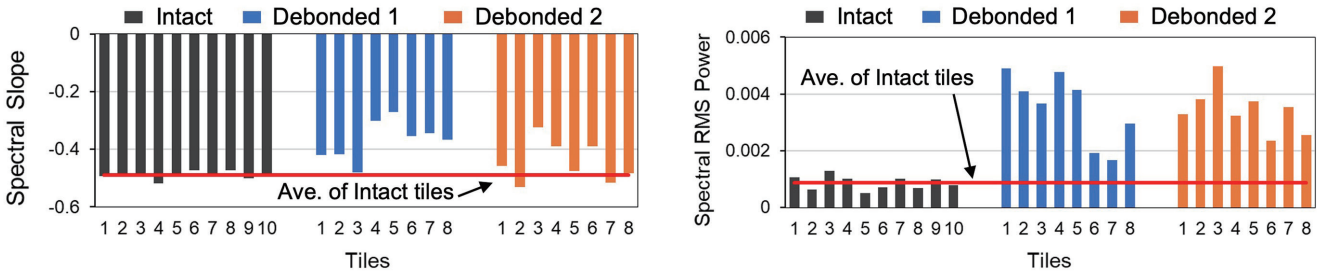
(h) Spectral Kurtosis



(i) Peak Frequency



(j) Peak Power(log scale)



(k) Spectral Slope (l) Spectral RMS Power

Fig. 8. Results of frequency-domain acoustic features

not imply that the indicator is ineffective. Since the calculation of Decay Time depends on the selected envelope threshold (currently set at 10% of the peak amplitude), modifying this threshold or employing an energy-integrated decay approach may yield clearer distinctions between bonding conditions. Therefore, this feature still has potential for optimization and improvement. In addition, as shown in Fig. 4, the envelope decay typically exhibits a rapid initial drop followed by a more gradual attenuation. Focusing on this rapid decay stage, for example by increasing the threshold level to capture this portion of the envelope, may help reveal differences between bonding conditions more clearly. However, this possibility requires further experimental verification and is left for future investigation.

4.4 Frequency-Domain Feature Results

As shown in Fig. 8, the frequency-domain features exhibit distinct differences among the three bonding conditions. The Spectral Centroid (Fig. 8 (a)) of the Intact samples is approximately two times higher than that of the debonded ones, indicating a greater proportion of high-frequency energy. For Spectral Bandwidth (Fig. 8 (b)), Debonded 1 shows the smallest value, suggesting a more concentrated frequency spectrum. Although this indicator has limited discriminative ability between "intact vs. debonded," it is useful for distinguishing between Debonded 1 and Debonded 2 after the general classification of bonding integrity is completed. The Spectral Roll-off Frequency (85%) (Fig. 8 (c)) of the intact samples is typically distributed around 9–12 kHz, while that of the debonded samples appears mainly in the 5–8 kHz range. This shift toward lower frequencies indicates changes in the dynamic response of the tile–substrate system caused by interfacial debonding. In addition to stiffness reduction, debonding may modify local boundary conditions and vibration modes, leading to a redistribution of spectral energy toward lower frequencies, as discussed in studies on interfacial defect detection using acoustic responses [23].

Regarding spectral complexity, the Spectral Entropy (Fig. 8 (d)) of intact samples is generally higher, whereas that of the debonded samples is lower. This indicates that, due to interface reflections and local resonances, acoustic energy in the debonded regions becomes concentrated in a limited number of frequency components, resulting in a more ordered spectral structure. Conversely, the intact tiles contain richer high-frequency components and more dispersed energy distribution. The Spectral Flatness (Fig. 8 (e)) exhibits an overall trend of Debonded 1 < Intact < Debonded 2 (partially overlapping samples). Debonded 1 still shows strong energy concentration, whereas Debonded 2 tends toward a "noise-like" distribution, indicating stronger reflection and scattering and a higher degree of randomness. Therefore, Spectral Flatness can serve as an auxiliary indicator to further differentiate the two types of debonding. The High-Frequency Energy Ratio (Fig. 8 (f)) is significantly higher for intact tiles, showing that debonding suppresses high-frequency transmission. It should be noted that the discriminative power of this feature could be improved by adjusting the high-frequency threshold (3000 Hz in this study).

In terms of spectral shape asymmetry and peak sharpness, the Spectral Skewness and Spectral Kurtosis (Fig. 8 (g) and Fig. 8 (h)) of the debonded samples are generally higher than those of the intact tiles. An increase in skewness indicates a shift of energy distribution toward the low-frequency range, while a higher kurtosis implies the presence of sharper dominant peaks and more concentrated local energy — consistent with the mechanism of low-frequency resonance generated by debonded cavities.

For the overall intensity and trend-related indicators, both the Peak Power (Fig. 8 (j)) and the spectral RMS power (Fig. 8 (l)) are markedly higher in the debonded samples (approximately 50 times and 4 times higher on average, respectively), suggesting that interface reflections are enhanced and overall acoustic energy increases. Combined with the downward shift of the Peak Frequency (Fig. 8 (i)), this implies that energy is further amplified in the low-frequency range, exhibiting a typical cavity-resonance effect. However, regarding the Spectral Slope (Fig. 8 (k)), the intact samples show more stable and concentrated slope values, indicating a relatively balanced energy distribution. Nevertheless, no clear trend is observed between intact and debonded tiles; thus, this feature is not recommended as a primary indicator for evaluating tile bonding integrity.

4.5 Statistical Analysis and Recommended Use of Feature Indicators

To provide an overall evaluation of the differences in the extracted acoustic features under different damage conditions, statistical analyses were conducted for the 22 feature indicators. Their statistical significance and recommended usage were summarized under two application situations. The two situations are defined as follows:

Situation A: comparison between intact tiles (Intact) and debonded tiles (Debonded 1 and Debonded 2 combined), which corresponds to detecting whether tile debonding occurs; Situation B: comparison between Debonded 1 and Debonded 2, which corresponds to distinguishing different types of debonding interfaces.

For the above two situations, two-sample statistical tests were performed to calculate the p-values for each feature indicator to determine whether the differences between groups were statistically significant. The p-value represents the probability of observing the current or more extreme differences if the null hypothesis (i.e., no difference between groups) is true. A smaller p-value indicates a lower probability that the observed differences are caused by random variation. In this study, $p < 0.05$ was adopted as the criterion for statistical significance. It should be noted that statistical significance only indicates that a difference exists between groups from a statistical perspective, and does not necessarily imply strong discriminative capability of the feature.

Table 1 summarizes the statistical analysis results of all feature indicators under the two situations. The results show that Fluctuation does not exhibit statistically significant differences in either situation, suggesting that the variation of this indicator among groups is likely caused by random factors. Therefore, it is not recommended as an evaluation indicator under the conditions of this study. In addition, strong correlations were observed among several feature indicators. For example, energy-related indicators (Maximum Absolute Value, RMS, and Energy) exhibit high correlations with each other, indicating that they reflect similar signal characteristics to some extent. Similarly, several frequency-location indicators, including Spectral Centroid, Spectral Roll-off, High-Frequency Energy Ratio, and Peak Frequency, also show strong correlations. In addition, Peak Power and Spectral RMS Power are strongly correlated, suggesting that both represent similar spectral power characteristics. Therefore, in practical applications, representative indicators can be selected according to specific requirements to reduce feature redundancy.

From a physical perspective, the different categories of indicators reflect distinct aspects of the hammering response. Energy- and amplitude-based parameters mainly describe the overall vibration intensity induced by impact excitation. Envelope-based parameters characterize the temporal decay behavior of the response and therefore reflect the persistence of structural resonance. In contrast, frequency-domain features capture the redistribution of spectral energy associated with changes in boundary conditions and vibration modes caused by interfacial debonding.

Finally, based on the statistical analysis results and the feature distributions shown in Fig. 6–8, the recommended usage situations of each feature indicator were summarized and presented in Table 1. These results provide references for subsequent feature selection and the design of detection methods.

5. Conclusions

This study conducted a systematic acoustic feature analysis of hammering signals to evaluate the bonding condition of exterior wall tiles. By comparing three typical states—Intact, tile–mortar interface debonding (Debonded 1), and mortar–substrate interface debonding (Debonded 2)—a total of 22 feature indicators were extracted from the time domain, envelope domain, and frequency domain to assess their effectiveness in identifying different bonding conditions. The main conclusions are summarized as follows:

(1) Several time-domain and envelope-based indicators can effectively distinguish intact and debonded conditions. Signals from debonded regions generally exhibit higher overall energy and slower attenuation, indicating stronger and more persistent vibration responses.

(2) Frequency-domain features reflect changes in the dynamic characteristics of the system. In debonded samples, the Spectral Centroid and Roll-off frequency shift toward lower frequencies, while the Peak Power and Spectral RMS Power increase significantly, suggesting reduced effective stiffness or changes in local boundary conditions and vibration modes, together with enhanced interface reflections and resonance effects.

(3) Statistical analysis indicates that most features show significant differences between intact and debonded conditions. Among them, RMS, Envelope Area, Spectral Centroid, Spectral Kurtosis, and Peak Power demonstrate relatively strong stability and discriminative capability and can therefore serve as promising input features for automated detection systems.

Overall, the multi-dimensional acoustic feature framework proposed in this study provides a comprehensive characterization of changes in tile bonding conditions and offers a quantitative basis for hammering-based inspection. However, it should be noted that the number of samples analyzed in this study was relatively limited, which may affect the statistical robustness of the results.

Future work will focus on expanding the dataset and validating these key features using machine learning models and exploring their integration into automated inspection systems such as mobile or robotic inspection platforms for intelligent detection of exterior tile debonding. Additional validation using complementary NDT methods (e.g., ultrasonic testing and infrared thermography) and pull-off tests will also be considered.

Table 1. Statistical analysis and recommended usage of acoustic feature indicators.

No.	Feature	p-value[*1] in situation A[*2]	p-value in situation B[*3]	Statistical significance[*4] (Situation)	Strong correlation with indicator No. ($r \geqslant 0.90$)[*5]	Recommended use[*6]
1	Maximum Absolute Value	<0.001	0.163	A	No.2 (r = 0.92) No.3 (r = 0.91)	A
2	Root Mean Square (RMS)	<0.001	0.868	A	No.1 (r = 0.92) No.3 (r = 0.96)	A
3	Energy	<0.001	0.719	A	No.1 (r = 0.91) No.2 (r = 0.96)	A
4	Crest Factor	<0.001	0.005	A and B		A
5	Decay Coefficient	<0.001	0.279	A		A
6	Envelope Area	<0.001	0.840	A		A
7	Early Energy Ratio	<0.001	0.950	A		A
8	Fluctuation Coefficient	0.382	0.116	Not significant		Not recommended
9	Temporal Centroid	<0.001	0.010	A and B		Discriminative capability may improve by reducing the calculation window (0–100 ms in this study)
10	Decay Time	0.002	0.445	A		Increasing the threshold (10% of the peak value in this study) may improve its discriminative capability
11	Spectral Centroid	<0.001	0.003	A and B	No.13 (r = 0.92) No.16 (r = 0.90) No.19 (r = 0.90)	A
12	Spectral Bandwidth	0.003	0.010	A and B		B
13	Spectral Roll-off	<0.001	0.010	A and B	No.11 (r = 0.92) No.16 (r = 0.92) No.19 (r = 0.91)	B
14	Spectral Entropy	<0.001	0.008	A and B		A
15	Spectral Flatness	0.091	0.048	B		B
16	High-Frequency Energy Ratio	<0.001	0.041	A and B	No.11 (r = 0.90) No.13 (r = 0.92) No.19 (r = 0.91)	A
17	Spectral Skewness	<0.001	0.839	A		A
18	Spectral Kurtosis	<0.001	0.916	A		A
19	Peak Frequency	<0.001	0.580	A	No.11 (r = 0.90) No.13 (r = 0.91) No.16 (r = 0.91)	A
20	Peak Power	0.003	0.318	A	No.22 (r = 0.92)	A
21	Spectral Slope	<0.001	0.992	A		Not recommended
22	Spectral RMS Power	<0.001	0.658	A	No.20 (r = 0.92)	A

*1: p-value represents the probability that the observed differences between groups occur by random chance under the null hypothesis. A significance level of $p < 0.05$ was adopted in this study.

*2: Situation A refers to the comparison between intact tiles and debonded tiles (Debonded 1 and Debonded 2 combined), which corresponds to the task of detecting the presence of debonding.

*3: Situation B refers to the comparison between Debonded 1 and Debonded 2, which corresponds to distinguishing between different debonding interface conditions.

*4: Statistical significance indicates that the difference in feature values between the compared groups is statistically significant. However, statistical significance does not necessarily imply strong discriminative capability.

*5: Highly correlated indicators may be interchangeable and used to reduce feature redundancy.

*6: Recommended use indicates the suggested application of each feature based on the observed feature distributions and comparisons presented in Fig. 6–8.

Acknowledgement

This work was supported by THE MAEDA ENGINEERING FOUNDATION and JSPS KAKENHI Grant Number JP23K26246.

References

[1] Yagi, S., Teramoto, A., Yeow, T., Seike, T., Kusunoki, K., and Nakamura, I. (2022) "Validating Resilient Detailing of Japanese Ceilings, Windows, and Wall Tiles Using an E-defense Shake-table Test," Journal of Earthquake Engineering, 26(16), pp. 8525~8551. https://doi.org/10.1080/13632469.2021.1988764.

[2] Cao, M.-T. (2023) "Drone-assisted segmentation of tile peeling on building façades using a deep learning model," Journal of Building Engineering, 80, 108063. https://doi.org/10.1016/j.jobe.2023.108063.

[3] Chew, M. Y. L. (2023) "Façade inspection for falling objects from tall buildings in Singapore," International Journal of Building Pathology and Adaptation, 41(6), pp. 162~183. https://doi.org/10.1108/IJBPA-10-2020-0087.

[4] Luk, B. L., Liu, L. K. P., and Collie, A. A. (2007) Climbing Service Robots for Improving Safety in Building Maintenance Industry, Bioinspiration and Robotics Walking and Climbing Robots, Maki K. Habib, Ed., I-Tech Education and Publishing, China.

[5] Luk, B. L., Liu, K. P., Tong, F., and Man, K. F. (2010) "Impact-acoustics inspection of tile-wall bonding integrity via wavelet transform and hidden Markov models," Journal of Sound and Vibration, 329(10), pp. 1954~1967. https://doi.org/10.1016/j.jsv.2009.11.038.

[6] Inoue, F., Doi, S., Okada, T., and Ohta, Y. (2009) "Development of Automated Inspection Robot and Diagnosis Method for Tile Wall Separation by Wavelet Analysis," 26th International Symposium on Automation and Robotics in Construction (ISARC 2009), pp. 379~388, Austin, USA.

[7] Kondo, T., and Kitagaki, R. (2014) "Report of the Research Committee on the Reliability Evaluation of an Automated Tile Debonding Diagnosis System for Building Facades (in Japanese)," Japan Society for Finishing Technnology, 26(155), pp. 36~39. https://doi.org/https://doi.org/10.14820/finexjournal.26.155_36.

[8] Seki, K., Iwasa, K., and Tsutsumi, H. (2022) "A study on the differences in the judgment results of sound hammering inspection in the diagnosis of RC exterior walls and its improvement by new technology (in Japanese)," Infrastructure Maintenance Practices, 1, pp. 394~401. https://doi.org/https://doi.org/10.11532/jsceim.1.1_394.

[9] Sonoda, Y., Nakayama, A., and Miyoshi, A. (2008) "A fundamental study on diagnostic mechanism of the rotary hammering test by acoustic analysis (in Japanese)," Journal of Structural Engineering, Japan Society of Civil Engineers (JSCE), 54, pp. 599~606.

[10] Lin, T.-H., Chiang, P.-C., and Putranto, A. (2024) "Multispecies hybrid bioinspired climbing robot for wall tile inspection," Automation in Construction, 164, 105446. https://doi.org/10.1016/j.autcon.2024.105446.

[11] Zhao, Y., Chen, Y., and Ye, L. (2023) "A non-contact inspection method of tile debonding using tuned acoustic wave and laser doppler vibrometer," Journal of Sound and Vibration, 564, 117875. https://doi.org/10.1016/j.jsv.2023.117875.

[12] Fujii, H., Yamashita, A., and Asama, H. (2015) "Improvement of Environ-mental Adaptivity of Defect Detector for Hammering Test Using Boosting Algorithm," 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 6507~6514, Hamburg, Germany.

[13] Soeta, T., and Mikami, T. (2016) "Basic study into a diagnostic system for exterior tile debonding (in Japanese)," Journal of Structural and Construction Engineering (Transactions of AIJ), 81(729), pp. 1779~1787. https://doi.org/10.3130/aijs.81.1779.

[14] Tong, F., Tso, S. K., and Xu, X. M. (2006) "Tile-wall bonding integrity inspection based on time-domain features of impact acoustics," Sensors and Actuators A: Physical, 132(2), pp. 557~566. https://doi.org/10.1016/j.sna.2005.12.035.

[15] Yamamoto, J., Ito, Y., Tanaka, D., and Kawabe, S. (2024) "The Influence of the Spring Constant and Free Length of a Small Solenoid Tapping Device on the Delamination Diagnosis of Exterior Wall Tile Finishing (in Japanese)," Proceedings of the Japan Concrete Institute, 46(1), pp. 1837~1842, Mastuyama, Ehime, Japan.

[16] Nagai, H., Tanaka, D., Ito, Y., and Kawabe, S. (2024), "The Delamination Diagnosis of Exterior Wall Tile Finishing Using Wavelet Analysis of Tapping Sounds by a Spatula Mounted on a Wall-Contact Drone (in Japanese)," Proceedings of the Japan Concrete Institute, 46(1), pp. 1849~1854, Mastuyama, Ehime, Japan.

[17] Inoue, F., Doi, S., Okada, T., and Ohta, Y. (2010) "Study on Automated Inspection Robot and Quantitative Detection of Outer Tile Wall Exfoliation by Wavelet Analysis," International Conference on Control, Automation and Systems 2010, pp. 994~999, Gyeonggi-do, Korea.

[18] Tanaka, D., Ito, Y., Kawabe, S., and Nagai, H. (2024) "The Study of a Delamination Diagnosis Method for Exterior Wall Tile Finishing Using Reverberation Time and Wavelet Coefficients through Wavelet Analysis (in Japanese)," Proceedings of the Japan Concrete Institute, 46(1), pp. 1819~1824, Mastuyama, Ehime, Japan.

[19] Emoto, H., Baba, Y., Asano, H., and Nagase, Y. (2020) "Comparision of AI method on hammering sounds at concrete bridge (in Japanese)," AI and Data Science Papers, 1(J1), pp. 514~521. https://doi.org/https://doi.org/10.11532/jsceiii.1.J1_514.

[20] Kitamura, Y., and Takayama, J. (2021) "Estimation of

spalling properties of wall surface based on time frequency analysis of hammering sound," Measurement: Sensors, 18, 100182. https://doi.org/10.1016/j.measen.2021.100182.

[21] Peeters, G. A large set of audio features for sound description (similarity and classification) in the CUIDADO project. IRCAM Technical Report, 2004.

[22] Tzanetakis, G., and Cook, P. (2002). "Musical genre classification of audio signals, " IEEE Transactions on Speech and Audio Processing, 10(5), pp. 293~302. https://doi.org/10.1109/TSA.2002.800560.

[23] Liu, H., Lin, X., Chen, Z., Liu, Y., Zhang, H., Wu, Y., & Lu, K. (2025). "Non-destructive inspection of interfacial debonding in concrete-filled steel tube bridges," Nondestructive Testing and Evaluation, 40(11), pp. 5197~5215. https://doi.org/10.1080/10589759.2024.2437679.