https://doi.org/10.1016/j.ymssp.2013.10.017

Acousto-ultrasonics-based Fatigue Damage Characterization: *Linear versus Nonlinear Signal Features*

Zhongqing SU [a,b,†], Chao ZHOU [c], Ming HONG [a,b], Li CHENG [a,b], Qiang WANG [d], Xinlin QING [e]

[a] The Hong Kong Polytechnic University Shenzhen Research Institute
Shenzhen 518057, PR China

[b] Department of Mechanical Engineering
The Hong Kong Polytechnic University, Kowloon, Hong Kong

[c] School of Mechanical and Electric Engineering
Guangzhou University, Guangzhou, Guangdong Province, 510006, PR China

[d] College of Automation
Nanjing University of Posts and Telecommunications, Nanjing, PR China

[e] Division of Aviation Health and Safety Management
Beijing Aeronautical Science and Technology Research Institute of COMAC
Beijing, 100083, PR China

Submitted to *Mechanical Systems and Signal Processing*

(initially submitted on 9 May 2013; revised and re-submitted on 14 June 2013)

[†] Corresponding author. Tel.: +852 2766 7818
Email address: mmsu@polyu.edu.hk (Z. Su).

Abstract

Engineering structures are prone to fatigue damage over service lifespan, entailing early detection and continuous monitoring of the fatigue damage from its initiation through growth. A hybrid approach for characterizing fatigue damage was developed, using two genres of damage indices constructed based on the linear and the nonlinear features of acousto-ultrasonic waves, respectively. The feasibility, precision and practicability of using linear and nonlinear signal features, for quantitatively evaluating multiple barely visible fatigue cracks in a metallic structure, was compared. Miniaturized piezoelectric elements were networked to actively generate and acquire acousto-ultrasonic waves. The active sensing, in conjunction with a diagnostic imaging algorithm, enabled quantitative evaluation of fatigue damage and facilitated embeddable health monitoring. Results unveiled that the nonlinear features of acousto-ultrasonic waves outperform their linear counterparts in terms of the detectability. Despite the deficiency in perceiving small-scale damage and the possibility of conveying false alarms, linear features show advantages in noise tolerance and therefore superior practicability. The comparison has consequently motivated an amalgamation of linear and nonlinear features of acousto-ultrasonic waves, targeting the prediction of multi-scale damage ranging from microscopic fatigue cracks to macroscopic gross damage.

***Keywords*:** signal processing; fatigue damage characterization; acousto-ultrasonics; nonlinear signal features; linear signal features; piezoelectric sensor network; structural health monitoring

1. Introduction

Acousto-ultrasonics, a coalescence of ultrasonic characterization and acoustic-emission, is one of the prevailing tools to develop non-destructive evaluation (NDE) and structural health monitoring (SHM) techniques [1–4]. Of particular interest in acousto-ultrasonics is the Lamb waves (the modality of acousto-ultrasonic (AU) disturbance guided by a thin sheet-like structure) in the ultrasonic regime. Inherently possessing appealing features including strong penetration, fast propagation, omnidirectional dissemination and high sensitivity to damage, Lamb-wave-based acousto-ultrasonics has been deployed in a diversity of fashions, showing demonstrated compromise among resolution, detectability, practicality, and cost [5–12]. The majority of such techniques are based on exploring changes in the damage-scattered AU waves, which can be documented in time domain signals in the form of amplitude alteration and/or phase deviation (in comparison with baseline signals), typified by the delay in time-of-flight (ToF) [12–16], wave reflection/transmission [17–20], energy dissipation [21,22] and mode conversion [23,24]. These signal features, for example the delay in ToF, show, to some extent, linear correlation with damage parameters such as the location, and are therefore colloquially referred to as *linear features* in this study.

On the other hand, there has been a consistent effort to reap the nonlinear features extracted from the damage-scattered AU waves to characterize material degradation [25] or structural damage [26–34]. Now on the verge of maturity for practical applications, the detection using nonlinear signal features is based on such a premise that AU waves, when propagating in an elastic medium, can be distorted by the inherent nonlinearity of the medium, resulting in an energy shift from the excitation to other frequency bands and generating nonlinear features such as high-order harmonics (contrastively called *nonlinear* *features* in what follows); upon occurrence of damage, micro-structures of the medium are altered, and the plastic zone in the vicinity of the damage incurs nonlinearities of AU waves. In addition, when AU waves traverse crack-like damage, the "breathing" motion pattern of the crack interface, under cyclic loads, creates localized nonlinear behaviors and introduces additional nonlinearity (generally called *contacting acoustic nonlinearity* (CAN) [26]). The nonlinear features which are often exploited by the approaches in this category include second-[27–30] or sub-[31] harmonics, mixed frequency responses [32] (*e.g.*, nonlinear wave modulation spectroscopy), shift of resonance frequency [33] (*e.g.*, nonlinear resonant ultrasound spectroscopy), dual frequency mixing [34], to name a few, as surveyed comprehensively elsewhere [35].

Yet, real-world structural damage often initiates from fatigue damage at imperceptible levels. Under cyclic loads the fatigue damage accumulates as the formation of dislocation monopoles, followed by dislocation loops and dipoles and subsequent dislocation veins and persistent slip bands. Fatigue cracks at the scale of few millimetres are then nucleated to microcracks, which can deteriorate and eventually coalesce to form macrocracks [36]. Under repetitious loads, the macrocracks can further grow and develop to a critical level at an alarming rate without sufficient warning, impacting detrimental effects on structural integrity and potentially resulting in catastrophic consequences. Early perception of small-scale fatigue damage has therefore become a cardinal measure to warrant the reliability, integrity and durability of ageing engineering structures, although it is a highly challenging task due to the small scales of the fatigue damage.

Both NDE and SHM techniques for characterizing fatigue damage, using either linear or nonlinear features of AU waves are in a good supply with diverse deployments, albeit the effectiveness and practicability of individual approaches are somewhat debatable. In the present study, a hybrid approach, using linear features (*i.e.*, delay in time-of-flight, and dissipation of wave energy) and nonlinear features (*i.e.*, second harmonic generation) extracted from AU wave signals, was developed. Two genres of damage indices were constructed, and respectively employed to evaluate barely visible fatigue cracks near rivet holes in a metallic structure. The feasibility, precision and practicability of using linear and nonlinear features were discussed comparatively. Miniaturized lead zirconate titanate (PZT) elements were networked and affixed to the structure, for generating and acquiring AU waves, which is deemed a critical step towards automatic and embeddable SHM.

2. Linear features for damage characterization

The ToF and wave energy are two sorts of most representative linear features which can be extracted from a captured AU wave signal [37]. In this study, these features acquired with a PZT sensor network in terms of pulse-echo and pitch-catch configurations are respectively associated with different damage parameters, for establishing linear damage indices (*DI*s).

2.1. ToF-based DI

ToF, the time spent for a wave packet to travel a certain distance, correlates the damage position with regard to the actuator and sensor in a sensor network (assuming the network comprises N PZT elements), according to

$$\left(\frac{\sqrt{(x_d-x_i)^2+(y_d-y_i)^2}}{V_{incident}}+\frac{\sqrt{(x_d-x_j)^2+(y_d-y_j)^2}}{V_{damage-scattered}}\right)-\frac{\sqrt{(x_j-x_i)^2+(y_j-y_i)^2}}{V_{incident}}=\Delta t_{i-j}\,,$$

$$(\,i, j = 1, 2, \ldots N,\; i \neq j) \qquad \textbf{(1)}$$

where $V_{incident}$ and $V_{damage-scattered}$ are the group velocities of the probing and the damage-scattered AU wave packets, respectively; Δt_{i-j} the difference between (i) the ToF for the probing AU wave to propagate from actuator S_i at (x_i, y_i) to damage at (x_d, y_d) and then to sensor S_j at (x_j, y_j), and (ii) the ToF for the probing AU wave to propagate from S_i to S_j directly, as illustrated in Fig. 1. Equation (1) mathematically depicts an ellipse-like locus, indicating all possible damage locations, perceived by sensing path $S_i - S_j$.

Inherently linking the damage location to the position of a known sensing path, ToF-based signal features can be used to define a *DI* with the assistance of a probabilistic imaging algorithm (PIA) [18,22,23]. The PIA differentiates itself from traditional damage imaging techniques such as tomography, taking advantage of its unique traits including in particular the use of an active sensor network with a much sparse transducer configuration instead of a dense network in tomography, and the adoption of a fast image reconstruction algorithm instead of computationally-expensive tomography. With PIA, the inspection region of the plate is meshed virtually, and projected to an image with each image pixel corresponding exclusively to a spatial point in the inspection region. The probability of damage presence at each spatial point is calibrated by the value borne by its corresponding pixel in the image (called *field value* hereinafter), in terms of the Euclid distance between a pixel to each locus defined by Eq. (1), as

$$F(z_i)\big|_{(x_m, y_n)} = \int_{-\infty}^{z_{ij}} f(z) \cdot dz\,, \tag{2}$$

where $f(z) = \dfrac{1}{\sigma_{ij}\sqrt{2\pi}} \exp[-\dfrac{z^2}{2\sigma_{ij}^2}]$, a Gaussian distribution function representing the probability density of damage presence at pixel node (x_m, y_n) (m = 1, 2, …, L; n = 1, 2, …, K, given the inspection region is rectangular and can be meshed using $L \times K$ nodes), established by sensing path $S_i - S_j$. z_{ij} is the Euclid distance between pixel (x_m, y_n) and the locus created by $S_i - S_j$, and σ_{ij} the standard variance. Equation (2) implies that the pixel right on a particular locus created by a sensing path has the highest probability (100%) of damage presence (from the perspective of that sensing path), while for other pixels the further the distance to this locus the lower the probability of damage is present at those pixels. Residing on Eq. (2), a *DI* is defined at each pixel for each sensing path in the sensor network, and for instance the one for path $S_i - S_j$, denoted by $DI_i(x_m, y_n)$, reads

$$DI_i(x_m, y_n)\Big|_{linear-ToF} = 1 - \Big[F(z_i)\Big|_{(x_m, y_n)} - F(-z_i)\Big|_{(x_m, y_n)} \Big]. \quad \textbf{(3)}$$

In the above, "*i*" and "*linear-ToF*" in the subscripts signify that the *DI* is defined at pixel (x_m, y_n) by the sensing path with S_i as the actuator, and it is based upon ToF-related linear signal features. The pixels with remarkably high field values are expected to highlight and further shape a damaged zone in the projected image, providing quantitative and detailed depiction about the damage (*e.g.*, size and orientation).

2.2. Energy-based DI

The discrepancy in damage (*e.g.*, different distances to a sensing path, or different shapes, severities and orientations [37]) can result in distinct magnitudes of damage-scattered AU wave energy. Therefore, deviation of the wave energy, determined from an AU signal captured from the structure under current inspection (called *current signal*), with regard to its counterpart captured from a pristine benchmark (*baseline signal*), can be employed to develop a *DI*, in terms of the correlation between the current and baseline signals, as

$$DI(i,j)\Big|_{linear-energy} = 1 - |\rho_{XY}(i,j)| = 1 - \left| \frac{\sum_{k=1}^{p} \left\{ HT\left[x_k(i,j)\right] - \eta_X(i,j) \right\} \left\{ HT\left[y_k(i,j)\right] - \eta_Y(i,j) \right\}}{\sqrt{\sum_{k=1}^{p} \left\{ HT\left[x_k(i,j)\right] - \eta_X(i,j) \right\}^2} \cdot \sqrt{\sum_{k=1}^{p} \left\{ HT\left[y_k(i,j)\right] - \eta_Y(i,j) \right\}^2}} \right|. \quad \textbf{(4)}$$

Similar to Eq. (3), "*ij*" and "*linear-energy*" allude to that the *DI* is defined at pixel (x_m, y_n) by $S_i - S_j$, based on energy-associated linear signal features. $\rho_{XY}(i, j)$ is the correlation coefficient between the current signal $X = \{x_1, x_2, ..., x_p\}$ acquired by $S_i - S_j$ and its corresponding baseline signal $Y = \{y_1, y_2, ..., y_p\}$. η is the signal mean and "*HT*" stands for Hilbert transform-processed signal. The use of Hilbert transform is aimed at producing an explicit envelope of the AU wave energy distribution and improving the recognisability of damage-scattered energy. The greater the similarity between $HT\left[x_k(i, j)\right]$ and $HT\left[y_k(i, j)\right]$, the closer to unity is $\rho_{XY}(i, j)$. A greater $\rho_{XY}(i, j)$ leads to a lower *DI* along path $S_i - S_j$, indicating a lower probability of damage existence near $S_i - S_j$; in contrast, in the case where damage is right on or close to $S_i - S_j$, $\rho_{XY}(i, j)$ becomes lower, resulting in a higher *DI*. Each sensing path, based on Eq. (4), contributes a probabilistic image in which the field value at each pixel is quantified in terms of $\rho_{XY}(i, j)$, indicating the probability of damage presence at the spatial points of the inspected structure correlated by that pixel.

2.3. Fusion of linear DIs

Each PZT element in the sensor network contributes two probabilistic images in terms of the ToF-based *DI* (Eq. (3)) and the energy-based *DI* (Eq. (4)), respectively. Called *source image*, each image is a prior perception on damage from the viewpoint of the sensing path creating the source image. With all sensing paths, a multitude of source images form a data pool, rendering such perceptions in plenty. In order to strengthen damage-related features (commonality in individual source images) and meanwhile dilute measurement noise/uncertainties (random information in individual source images), all source images are fused at the pixel level, leading to an ultimate image.

Note that during the fusion, the *DI* defined by Eq. (3) is given at each pixel, whereas the *DI* described by Eq. (4) is defined along a sensing path (*viz.*, all the pixels along a sensing path holding the same *DI*). Compatibility between two *DI*s must be reached provided the fusion is carried out at the pixel level. It can be seen that in Eq. (4), $DI(i,j)\big|_{linear-energy}$ is calculated when S_i serving as the actuator, and thus each actuator contributes $N-1$ probabilistic images using PIA. To achieve such compatibility (*i.e.*, a transform from a sensing path defined by Eq. (4) to a pixel), all these images are pre-aggregated, to create a source image for S_i at (x_m, y_n), denoted by $DI_i(x_m, y_n)\big|_{linear-energy}$, according to

$$DI_i(x_m, y_n)\big|_{linear-energy} = \frac{1}{(N-1)} \sum_{j=1}^{N(j \neq i)} DI(i,j)\big|_{linear-energy}, \quad \textbf{(5)}$$

where "*i*" in the subscript accentuates that the *DI* is now re-defined at pixel (x_m, y_n) for the sensing path with S_i as the actuator. Conclusively, each actuator in the sensor network ends up with a source image via ToF-based *DI* (Eq. (3)), and another source image via energy-based *DI* (Eq. (5)), which are then fused by

$$DI(x_m, y_n)\big|_{linear} = \frac{1}{N} \sum_{i=1}^{N} \left(DI_i(x_m, y_n)\big|_{linear-ToF} \cap DI_i(x_m, y_n)\big|_{linear-energy} \right), \quad \textbf{(6)}$$

where $DI(x_m, y_n)\big|_{linear}$ is the fused *DI* in the ultimate image, based on all the extracted linear features, reflecting the probability of damage presence at each pixel. In Eq. (6) an arithmetic fusion ('Σ') takes into account the prior perceptions from all source images and equally decentralizes individual contributions. But arithmetic fusion is anticipated to embrace ambient noise and measurement uncertainty as well. Thus, a conjunctive fusion ('$\cap$') multiplicatively processes source images to supplement the arithmetic fusion, with which a low field value at a pixel in any source image due to noise or uncertainty can lead to a significantly low likelihood of damage presence at that pixel in the ultimate image, effectively eliminating the measurement noise and uncertainties.

3. Nonlinear features for damage characterization

In parallel with the linear features, nonlinear features are extracted from the same AU wave signals to establish a nonlinear *DI*. In this connection, most existing efforts are of a nature of qualitative detection (capable only of indicating damage existence), and extension to quantitative and automatic SHM are fairly hampered due to the use of bulky ultrasonic probes. In this study, in conjunction with the use of the above active PZT sensor network, the developed nonlinear *DI* can be endowed with a capacity of characterizing fatigue damage quantitatively, facilitating embeddable SHM.

3.1. Theory

For an undamaged isotropic solid medium, two types of nonlinearity need to be addressed: the material nonlinearity and the geometric (or convective) nonlinearity. The former inherently originates from the nonlinear elastic properties of the medium (*viz.*, the lattice elasticity), while the latter from the mathematic transformation of wave motion equation from the Eulerian to the Lagrangian coordinate systems [38]. Both can be described, using the second-order nonlinear approximation, as

$$\sigma_{ij} = (C_{ijkl} + 1/2\, M_{ijklmn}\varepsilon_{mn})\varepsilon_{kl}\,, \qquad \textbf{(7)}$$

where σ_{ij} is the stress tensor; ε_{mn} and ε_{kl} the strain tensors. C_{ijkl} and those in a similar form in followings with different subscript ordering are the second-order elastic (SoE) tensors defined with Lamé's constants λ and μ; and M_{ijklmn} a tensor embracing the above two types of nonlinearity simultaneously, which reads

$$M_{ijklmn} = C_{ijklmn} + C_{ijln}\delta_{km} + C_{jnkl}\delta_{im} + C_{jlmn}\delta_{ik} \,, \tag{8a}$$

where

$$C_{ijklmn} = \frac{1}{2} A(\delta_{ik} I_{jlmn} + \delta_{il} I_{jkmn} + \delta_{jk} I_{ilmn} + \delta_{jl} I_{ikmn}) \; + \; 2B(\delta_{ij} I_{klmn} + \delta_{kl} I_{mnij} + \delta_{mn} I_{ijkl}) \; + \; 2C\delta_{ij}\delta_{kl}\delta_{mn} \,. \tag{8b}$$

In the above, δ_{km} and those in a similar form with different subscript ordering (δ_{im}, *etc.*) are the Kronecker deltas; I_{jlmn} and those in a similar form (I_{jkmn}, *etc.*) the fourth-order identity tensors; C_{ijklmn} the third-order elastic (ToE) tensor addressing material nonlinearity. *A*, *B*, and *C* are three ToF constants. The last three terms in Eq. (8a) all together reflect the geometric nonlinearity. In an extreme occasion that the second-order nonlinear term ($1/2\, M_{ijklmn}\varepsilon_{mn}$) eliminated, Eq. (7) reverts to the three-dimensional Hooke's Law for linear elasticity.

Without loss of generality, consider a one-dimensional medium for illustration, and Eq. (7) can be re-written, using a quadratic approach, as

$$\sigma = (E + E_2\varepsilon)\varepsilon \,, \tag{9}$$

where σ, ε, E and E_2 are the stress, strain, first-order (reflecting linear property) and second-order (reflecting nonlinear property) Young's moduli of the medium, respectively [39]. Combining Eqs. (7) – (9) yields

$$E_2 = -\frac{1}{2}\left(3E + 2A + 6B + 2C\right), \tag{10}$$

and further

$$\beta_g = \frac{E_2}{E} = -\frac{1}{2}(3 + \frac{2A + 6B + 2C}{E}), \tag{11}$$

where β_g is the ratio of two Young's Moduli. It can be seen that all the parameters in Eqs. (10) and (11) are pertaining to the SoE and ToE constants, and thus for an ideal material in its pristine status without any fatigue damage or plastic deformation, β_g is a constant at a given measurement distance, serving as an intrinsic material property accounting for the nonlinearity caused by the material's lattice anharmonicity.

The occurrence of fatigue damage brings about additional nonlinear sources in its vicinity. Taking this into account, a twofold coefficient, β, is introduced

$$\beta = \begin{cases} \beta_g & \text{(without fatigue damage)} \\ \beta_g + \beta_l, & \text{(with fatigue damage)} \end{cases} \quad \textbf{(12)}$$

where β_l is a localized nonlinearity coefficient addressing the nonlinearity contributed by the fatigue damage alone. The authors' previous study [39] has demonstrated that β_l plays a dominant role in the generation of nonlinearity in AU waves, much prominent than β_g.

To deploy β_l in an explicit modality, recall the governing equation for the above one-dimensional medium

$$\rho \frac{\partial^2 u(x,t)}{\partial t^2} = \frac{\partial \sigma}{\partial x}, \quad \textbf{(13)}$$

where ρ is the density of the medium, $u(x,t)$ the particle displacement at x along propagation direction at instant t (abbreviated as u). Using a perturbation theory [29,40], Eq. (12) can be solved, leading to

$$u = A_1 \cos(kx - \omega t) + A_2 \cos(2kx - 2\omega t), \quad \textbf{(14)}$$

where $A_2 = \frac{\hat{\beta}}{8} A_1^2 k^2 x$. ω is the angular frequency of the excitation and k the wavenumber, respectively; A_1 and A_2 the magnitudes of the probing wave mode (with a frequency of

ω, called *fundamental mode*) and its second harmonic wave mode (with a frequency of 2ω, called *second harmonic mode*), respectively. $\hat{\beta}$ denotes the acoustic nonlinearity parameter, which after rearrangement reads

$$\hat{\beta} = \frac{8}{k^2 x} \frac{A_2}{A_1^2}. \quad \textbf{(15)}$$

Previous studies [29,30] have demonstrated that an increase in β due to the presence of fatigue damage can be faithfully, if not totally, reflected by the increase in $\hat{\beta}$, meaning that the nonlinearity originated from the material itself is insignificant compared with that arising from the fatigue damage. Therefore, any singular increase in $\hat{\beta}$ is able to pinpoint the occurrence of fatigue damage. Based on this, the nonlinearities associated with the medium and the damage can be determined by probing A_1 and A_2 from a captured AU wave signal. To detect the fatigue damage, one is more interested in the change in $\hat{\beta}$ than its absolute value, and thus for a given wave propagation distance a relative acoustic nonlinearity parameter β' is further defined as

$$\beta' = \frac{A_2}{A_1^2}. \quad \textbf{(16)}$$

As seen, β' is proportional to β and addresses the essential nonlinearity of a captured AU wave signal subject to fatigue accumulation, therefore able to serve as a primary index for quantitative characterization of fatigue damage. Note that Eq. (16) is defined for a one-dimensional medium, while for Lamb waves in plates, such an index can be achieved by multiplying a scaling factor [30], because a medium has an unchanged scaling factor at a given measurement point regardless of the occurrence of fatigue damage.

3.2. β'-based DI

To develop a *DI* capitalizing on β', first, a correlation is established between (i) the relative distance from the fatigue damage to a particular sensing path in the sensor network (this relative distance is referred to as *measurement deviation* (MD) hereinafter) and (ii) the value of β' extracted from the AU wave signals acquired via that sensing path [29]. Fig. 2 exemplarily shows such a correlation for an aluminum plate with a thickness of 4.5 mm, where *MD* is normalized with regard to the wavelength of the fundamental mode (*MD/λ*), making it possible to extend the results to general circumstances at other excitation frequencies. It has been shown that the smaller *MD* is, the higher β' is, presenting approximately monotonous variation. In addition, such a correlation is observed to be insensitive to the difference in the length of a sensing path [29], which can be attributable to the fact that compared with the cumulative material nonlinearity along with wave propagation, the one incurred due to fatigue damage dominates the overall nonlinearity manifested in AU wave signals. It is also relevant to note that β' captured via a sensing path possesses high inertness to distant damage away from that path, implying a sensing path perceives the damage near it only. Such a trait makes it possible to identify multi-fatigue damage using such a nonlinear parameter.

Residing on β', a *DI* is constructed using the aforementioned PIA, defined at pixel (x_m, y_n) (denoted by $DI_i(x_m, y_n)\big|_{nonlinear-\beta'}$), as

$$DI_i(x_m, y_n)\big|_{nonlinear-\beta'} = \beta_i' \left[\frac{\varsigma - R_i(x,y)}{\varsigma - 1} \right]. \quad \textbf{(17)}$$

The subscripts "*i*" and "*nonlinear-* β'" stress that the index is defined for the sensing path with S_i being the actuator, and it is obtained upon β'-related nonlinear signal features. In Eq. (17), ς is a scaling parameter controlling the size of the effective distribution area, and

$R_{ij}(x,y)$ a weight to regulate the area of influence from the fatigue damage on a sensing path [22,37] which reads

$$R_i(x,y)=\begin{cases}\dfrac{\sqrt{(x_m-x_i)^2+(y_n-y_i)^2}+\sqrt{(x_m-x_j)^2+(y_n-y_j)^2}}{\sqrt{(x_i-x_j)^2+(y_i-y_j)^2}} & when\ \ r_i(x,y)<\varsigma \\ \varsigma. & when\ \ r_i(x,y)\geq\varsigma\end{cases} \quad \textbf{(18)}$$

With Eq. (18), each sensing path in the sensor network contributes a probabilistic image. Ideally, all the field values are low provided the inspection area is free of fatigue damage (practically it is not zero due to noise interference), while they are elevated pronouncedly at those pixels contained in the fatigue damage zone (subject to *MD*).

4. Linear *versus* nonlinear *DI*s

Both the linear and the nonlinear *DI*s were applied comparatively, to evaluate barely visible fatigue cracks in metallic plates.

4.1. Feasibility study

As a preliminary evaluation, two genres of *DI*s were first used to identify a mono-fatigue crack in an aluminum panel, to examine their respective effectiveness in evaluating fatigue cracks in a simple case.

4.1.1. Specimen preparation and measurement configuration

An aluminum plate ($484\times300\times2.2$ mm^3) was prepared as shown in Fig. 3. To introduce a fatigue crack in the plate, a sharp notch was machined at the center of the upper edge. The plate was fatigued under a sinusoidal tensile load with a magnitude of 4 kN at a frequency of 5 Hz using a digitally controlled fatigue testing machine (MTS 810). It took about

50,000 cycles to produce a fatigue crack (circa 5 mm in length originating from the notch tip). Four PZT wafers (nominal diameter: 6.9 mm, thickness: 0.5 mm each) were surface-mounted on the plate, denoted by PZT_i ($i = 1, 2, 3, 4$), and instrumented with a signal generation and acquisition system developed on a *VXI* platform [41]. Referring to Fig. 3 for respective coordinates, these four PZT wafers in principle provided $4 \times 3 = 12$ sensing paths. The probing fundamental mode, which was different for constructing the linear and the nonlinear *DI*s (to be detailed in the next sections), was generated in MATLAB® and downloaded to an arbitrary waveform generation unit (Agilent® E1441), and then amplified with a signal amplifier (US-TXP-3) to 80 $V_{p\text{-}p}$, which was then applied in turn on each PZT wafer; the signals sensed by the remaining three wafers were acquired with a signal digitizer (Agilent® E1438) at a sampling rate of 40 MHz.

4.1.2. Linear DI

To construct the ToF-based and energy-based linear *DI*s defined by Eqs. (3) and (5), respectively, five-cycle Hanning-windowed sinusoidal tone bursts at a central frequency of 300 kHz were excited. The selection of the current frequency facilitated generation of the fundamental symmetric wave mode, under which the wave signals were observed to feature the best signal recognizability for extracting linear signal features. As a representative example, the time domain signal acquired via sensing path $PZT_2 - PZT_3$ (a sensing path right traversing the fatigue damage) for the current state (with a fatigue crack) and its corresponding baseline signal from the pristine counterpart (notched but before the fatigue treatment) are compared in Fig. 4. The $\rho_{XY}(2,3)$ was calculated to be 0.9408 and consequently $DI(2,3)\big|_{linear-energy}$ be 0.0592 in terms of Eq. (4), indicating the discrepancy between the current and the baseline signals is minute. Meanwhile, $DI_2(x_m, y_n)\big|_{linear-ToF}$

along $PZT_2 – PZT_3$ also presents a low value (using Eq. (3)), because of the difficulty in identifying damage-scattered wave packet. A primary conclusion can thus be drawn that the linear *DI*, attempting to explore changes in the linear signal features such as ToF, energy attenuation, transmission and reflection might be ineffective to deal with small-scale fatigue damage because no phenomenal linear wave scattering can be observed in the captured AU wave signals.

4.1.3. Nonlinear DI

In order to explore an optimal excitation frequency at which the nonlinear features of AU waves upon interaction with fatigue damage can be prominent, Gaussian white noise was applied on PZT_2 as an input signal, and the frequency spectrum of the signal captured by PZT_3, obtained using fast Fourier transform (FFT), is exhibited in Fig. 5. A strong response can be observed at 380 kHz, which was therefore selected as the excitation frequency to modulate the five-cycle Hanning-windowed sinusoidal tone bursts. This frequency has proven effectiveness in generating AU wave signals with an enhanced signal-to-noise ratio, benefiting extraction of nonlinear signal features and improvement of accuracy of β' calculation. Note there was a slight difference in excitation frequency for constructing the linear and nonlinear *DI*s, which was aimed at achieving the highest signal-to-noise ratio and best signal recognizability, for extracting linear and nonlinear wave features, respectively.

For illustration, the signal spectra, acquired via path $PZT_2 – PZT_3$ under excitation of 380 kHz when the panel was in its pristine status and in fatigued status are displayed in Figs. 6(a) and (b), respectively, to observe that for the pristine status, the majority of the AU wave energy is concentrated near the excitation frequency, whereas for the fatigued status there is an obvious shift of probing energy from the excitation to other frequency bands, as evidenced by the occurrence of side lobes at 760 kHz (second harmonic) and even 1.14 MHz (third harmonic). These nonlinear features captured at twice the fundamental frequency can be extracted to construct nonlinear *DI* using Eq. (17).

4.2. Quantitative evaluation of fatigue cracks near rivet holes

With demonstrated feasibility, the proposed approach was then applied to characterize multi-fatigue cracks near rivet holes in an aluminum plate.

4.2.1. Specimen preparation and measurement configuration

An aluminum plate ($380 \times 400 \times 4.5$ mm^3) containing four through-thickness rivet holes for bolt connection (diameter: 10 mm each), as schematically shown in Fig. 7(a), was fatigued using the aforementioned fatigue processing. To accelerate initiation of fatigue cracks, two stress risers were inscribed at the edges of Hole 1 and 2, respectively. Fatiguing the plate after 500,000 cycles led to two hairline barely visible fatigue cracks, as displayed in Fig. 7(b), with one measuring 5 mm in length near Hole 1 and the other 3 mm near Hole 2. Upon completion of fatigue testing, a sensor network comprising eight circular PZT wafers (nominal diameter: 5 mm, thickness: 0.5 mm each) were surface-mounted on the fatigued plate, denoted by PZT_i ($i = 1, 2, \cdots, 8$), as seen in Fig. 7. All PZT wafers were instrumented with the signal generation and acquisition system introduced previously. The configured sensor network rendered $7 \times 8 = 56$ sensing paths.

4.2.2. Linear DI

Fig. 8 displays the probabilistic image using the PIA, upon fusing ToF-based *DI* (defined by Eq. (3)) and energy-based *DI* (defined by Eq. (5)) using the fusion algorithm described by Eq. (6). In the figure, the presence probability of fatigue damage is calibrated in greyscale, where the darker a pixel the greater the presence possibility of fatigue damage at that pixel it is.

The two fatigue cracks could not be identified in the image, and such a failure can be attributed to the small-scale of the fatigue cracks which was unable to produce phenomenal wave scattering (the wavelength of the probing wave is around 27 mm at the current excitation frequency, much greater than the major dimension of the fatigue crack), therefore failing to generate observable changes in ToF and wave energy. The difficulty in extracting linear features resulted in abundant pseudo prediction in the figure, under the interference from ambient noise. As the linear signal features are of the same order of ambient noise, the diagnostic results of the probabilistic image pessimistically exaggerate the possibility of damage occurrence, interpreting this observation.

It is noteworthy that both linear *DI*s defined by Eqs. (3) and (5) seek the difference between a current signal and a baseline signal. Based on such a philosophy, the rivet holes and fatigue crack imitators at the hole edges would not, in principle, be detected using the linear *DI*s, because they are the connatural geometric features of the sample in both intact and damaged statuses. Provided the pristine plate before the introduction of the rivet holes can be benchmarked, these gross damage cases can be identified using the two linear *DI*s, as reported in the authors' previous work [37,41]. It might be helpful to use the linear *DI*s if one increases the excitation frequency to reach smaller wavelengths that are comparable with the fatigue crack size. However, as AU waves are of dispersive nature, the multiple modes gradually appearing at higher frequencies would make it highly challenging to extract linear features. In contrast, such a barrier may not be a concern for the approach capitalizing on nonlinear *DI* as no time domain features will be explored.

4.2.3. Nonlinear DI

To achieve conspicuous generation and accumulation of the desired second harmonic for establishing nonlinear *DI*, the fundamental and second harmonic modes in Eqs. (14)-(16) should ideally satisfy a twofold prerequisite: (i) *synchronism*: both the phase and the group velocities of the fundamental mode match those of the second harmonic mode, respectively and concurrently; (ii) *non-zero power flux*: the fundamental mode is of the same type as the second harmonic mode (*e.g.*, both are either symmetric or anti-symmetric). This guarantees the shift of AU energy from the fundamental to the second harmonic modes with increasing propagation distance [28]. Fig. 9 shows the calculated (using DISPERSE®) dispersion curves of AU waves propagating in an aluminum plate, in which the modes S_1 (the first-order symmetric Lamb mode) at 3.57 MHz·mm and S_2 (the second-order symmetric Lamb mode) at 7.14 MHz·mm, as highlighted in the figures, meet the above prerequisite and form a synchronous pair. In addition, at 3.57 MHz·mm and 7.14 MHz·mm, the S_1 and S_2 modes respectively propagate at the same highest speed among all available modes, simplifying their isolation from other wave modes. Allowing for the thickness of the plate (4.5 mm), sixteen-cycle Hanning-windowed sinusoid tone bursts at a central frequency of 800 kHz were applied as the probing wave. The magnitude of S_1 at 3.57 MHz·mm (corresponding to A_1 in Eqs. (14) – (16)) and that of S_2 at 7.14 MHz·mm (corresponding to A_2) were calculated from signals captured via available sensing paths.

As a representative, the time domain signal captured via sensing path $PZT_2 – PZT_7$ is presented in Fig. 10, which was observed to be identical with its corresponding baseline signal prior to fatigue processing. The multimodal and dispersive natures of the AU waves embarrass proper recognition of the synchronous pair in the time domain, as interpreted previously that the small-scale of the fatigue damage would not incur noticeable wave scattering in the time domain. This has entailed meticulous signal processing to improve the signal recognizability. To this end, short time Fourier transform (STFT) was used to deploy the signal over a time-frequency domain, as shown in Fig. 11 for the signal in Fig. 10. In the time-frequency spectrum, the fundamental and second harmonic modes were extracted at 800 kHz and 1.6 MHz, respectively; both were then re-constructed to the time domain and are combined in Fig. 12, where A_1 and A_2 were determined to calculate β' using Eq. (16). It is relevant to note that the slight difference in the arrival time of two wave modes can be attributed to measurement noise and uncertainties. Subsequently, the nonlinear *DI* was constructed using Eq. (17), leading to a source image as shown in Fig. 13. This image, like Fig. 8, reflects the presence probability of fatigue damage at each pixel. Notably, Fig. 13 corroborates the conclusion previously drawn in Section 3.2 that β' captured via a sensing path possesses high inertness to distant damage (manifested as a very narrow dark area centralized along the sensing path in the source image), benefiting identification of multi-fatigue damage.

Repeating the above procedure for all available sensing paths in the network, 56 source images in total formed a probabilistic image pool, and fusion of these source images using an arithmetic mean algorithm with a threshold correction [29] yielded an ultimate image, in Fig. 14. The ultimate image highlights explicitly two regions near the edges of Hole 1 and Hole 2 with higher probability of fatigue damage occurrence, coinciding with reality; in contrast, regions away from the two fatigue cracks present much lower field values.

It is interesting to note that the highlighted regions with higher greyscale in Fig. 14 are greater than the actual sizes of the two fatigue cracks, which can be attributable to the fact that the plastic zone in the vicinity of the fatigue damage also increased β' and consequently the field values therein. It is also relevant to emphasize that the highlighted regions are corresponding to the two fatigue cracks initiated from the rivet holes, rather than the rivet holes themselves or the fatigue crack initiators at the hole edges, because, as explained previously, this approach explores the abnormal increase in β' due to a fatigue crack only, rather than the connatural material nonlinearity, geometric nonlinearity and gross damage (*e.g.*, a rivet hole in this study). Thus, once increase in β' is detected, it is predicted that fatigue damage exists. By the same token, the nonlinear *DI* failed to identify gross damage such as the rivet holes, because the rivet hole would not incur significant change in β' before and after the fatigue process. Both fatigue cracks are revealed in the ultimate image simultaneously, corroborating the effectiveness of the nonlinear *DI* in evaluating multiple fatigue cracks, a trait of the *DI* which is highly inert to distant damage away from the sensing path via which the *DI* is constructed.

5. Concluding remarks

It is significant but also challenging to detect fatigue cracks at a quantitative level. An hybrid approach, in conjunction with a probability-based diagnostic imaging algorithm, for characterizing fatigue damage was developed, capitalizing on two genres of *DI*s developed using linear and nonlinear features extracted from acousto-ultrasonic waves, respectively. Typical linear AU wave characteristics (*i.e.*, delay in ToF and damage-scattered wave energy) and nonlinear features (*i.e.*, second harmonic generation) were extracted from AU wave signals acquired by an active sensor network, and used, respectively, to construct different *DI*s. The use of the active sensor network enabled an extension of the traditional means for capturing nonlinear wave features to embeddable health monitoring, which however is at the expense of introducing complexity in extracting the nonlinearity of AU waves using PZT wafers such as the weak magnitudes of the nonlinear features, requesting deliberate selection of wave mode, excitation frequency, and signal processing tools (*e.g.*, STFT demonstrated in this study). Based on comparison of respective feasibility, precision and practicability when evaluating barely visible fatigue cracks in a metallic plate, it has been revealed that the nonlinear AU wave features have higher sensitivity than linear signal features and therefore superior detectability for small-scale fatigue damage, mainly due to the fact that under the modulation of traversing waves, fatigue cracks present nonlinear characteristics, which may not be strongly evidenced in the linear macroscopic changes of AU waves. The proposed nonlinear *DI* possesses high inertness to distant damage, making it possible to identify multi-fatigue damage. The detection can be at a quantitative level, including the co-presence of multi-cracks, and their individual locations and severities. The study has consequently motivated proper amalgamation of the linear and nonlinear features of AU waves, reaching a capacity of characterizing multi-scale damage ranging from microscopic fatigue cracks to macroscopic gross damage.

Acknowledgement

This project is supported by National Natural Science Foundation of China (Grant No. 51375414, 11272272 and 11202107). This project is also supported by the Hong Kong Research Grants Council via a General Research Fund (GRF) (No. 523313). Qiang Wang is grateful for the Doctoral Program of Higher Education (Grant No. 20113223120008) and Natural Science Foundation of Jiangsu Higher Education Institutions of China (Grant No. 11KJB130002).