CAAP Annual Report

Date of Report: *October 7, 2019*

Contract Number: *693JK318500010CAAP*

Prepared for: *U.S. DOT Pipeline and Hazardous Materials Safety Administration*

Project Title: Brain-Inspired Learning Framework to Bridging Information, Uncertainty and Human-Machine Decision-Making for Decoding Variance in Pipeline Computational Models

Prepared by: *North Dakota State University*

Contact Information: *Ms. Zi Zhang, PhD student, Email: zi.zhang@ndsu.edu@ndsu.edu, Phone: 701-231-7204; Mr. Matthew Pearson, M.S. student, Email: matthew.pearson@ndsu.edu, Phone: 701-231-7204; Dr. Zhibin Lin, Email: zhibin.lin@ndsu.edu, Phone: 717-231-7204*

For quarterly period ending: *October 7, 2019*

<u>Business and Activity Section</u>

(a) Generated Commitments

No changes to the existing agreement
Some purchase of steel plates and piezoelectric sensors

(b) Status Update of Past Quarter Activities

In the fourth report, the major work aimed to modify the simulated model, introduce data-driven method to analyze the received signals and laboratory test associated with varying mechanical damages, while numerical simulation enriched datasets for mechanical damage, as summarized below:

1. Background in the annual report

Large-scale oil/gas pipeline systems are lifelines for economic and social need. Similar to other transportation systems, large-scale networked onshore gas and liquid transmission pipelines have to face with the harsh environments and extreme events over time. As a result, these pipelines are susceptible to certain levels of degradation, corrosion, and damages due to aging, loads, and man-made disasters. Increasing research endeavors and the technological progress in recent years have been directed to the development and implementation of sensor technologies and associated information fusion to assess new and existing pipeline systems to improve structural integrity. However, the pipeline industry has to face with the great challenges in terms of the high level of variances. Thus, this study aimed to provide new solution for assisting understanding the variances.

2. Objectives and methodology

The objectives of this report were: (a) Collection of datasets (from both experimental and numerical studies) of mechanical damage/defects; (b) Understanding of sensitive features (e.g., damage type, orientation, and size) using machine learning;

2.1. Methodology

The statistical pattern recognition paradigm in SHM applications has recently been proposed, as shown in the flow chart in **Fig. 1**. The data mining process is classified into four steps [1,2]: (a) operational evaluation; (b) data acquisition; (c) feature extraction; and (d) statistical model development for feature classification. This framework displays the workflow from sensory data acquired from sensor systems to sensitive feature extraction. SVM learning algorithms are specifically designed for classification between damage and undamaged cases, where the radial basis function (RBF) kernel are herein chosen as the kernel function, as detailed below.

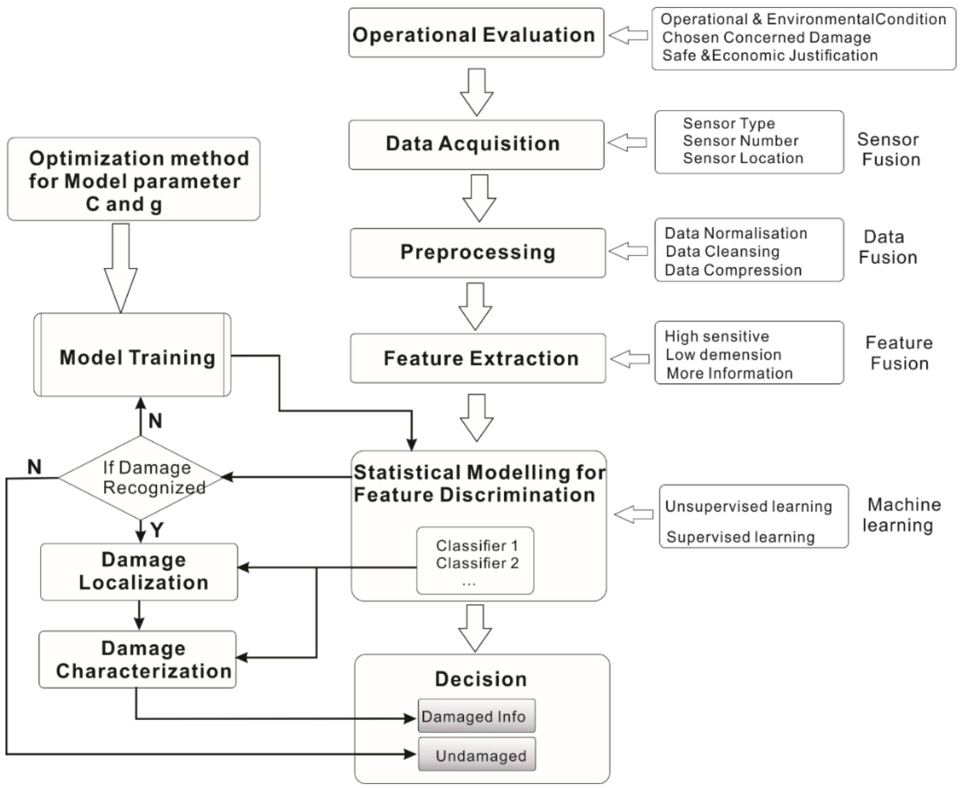


Fig. 1. Data mining process for identification [1, 2]

Thus, the detailed tasks as presented as follows included:

(a) Collection of datasets, using experimental and numerical plans through lamb wave;
(b) Data fusion
(c) Feature representation and classification.

2.2. Collection of datasets

Guided wave analysis has been proposed for damage detection in early 1990. Several forms of guided wave have been used in SHM, such as axial wave, flexural wave, shear wave, Rayleigh wave and lamb wave. Axial wave is longitudinal wave and shear wave has particle displacement perpendicular to the direction of propagation. Rayleigh wave is a surface wave, while lamb wave is a type of ultrasonic wave.

Therefore, this study aimed to collect a large number of datasets from both experimental and numerical investigations to enrich data types to simulate potential information with various variances experienced in pipeline systems in fields. The following sections were summarized from experimental and numerical standpoints, while the results and discussion were presented in detail in Section 3.

2.2.1 Data generated from experiment testing

This study attempted to collect data from experimental testing, which could provide certain level of uncertainty due to laboratory conditions. The experiment consists of generator, oscilloscope, Piezo actuators and a steel plate. The generator submits the voltage signal with different mode. Then actuator changes the voltage signal into mechanical signal. The wave propagates in the steel plate. When it arrives the edge or damage of the steel plate, the signal can be reflected and received by the second actuator. Next, the piezo actuator changes the wave into voltage signal. Different damage types were designed on the steel plate.

2.2.2 Data generated from computation modeling of lamb wave

Lamb wave is wildly used in non- destructive testing for damage detection. The location and severity of the damage can be detected by analyzing the changes of the lamb wave signal. Lamb wave excited in thin plate can appear different modes, symmetric mode (S mode) and anti-symmetric mode (A mode). Some research**Error! Reference source not found.** proved that the lowest frequency wave mode S_0 and A_0 has the sensitive to inspect the damages. S_0 mode is sensitively for internal damages in thin plate-like structure. While A_0 mode is sensitively for surface damages.

2.3.Data fusion

Lamb wave exhibits apparently non-stationary and nonlinear behavior. Time domain/frequency/time-frequency analyses are effective to track the change of a system and its nonlinear behavior and the conventional techniques are mostly encompassed by the Wavelet transform, short-time Fourier transform and Wigner-Ville distribution. These methods have their own limitation in noise sensitivity. Literature review shows that few attempts are made to address impacts of various feature extraction methods on structural condition assessment and damage detection. Thus, we discussed representative feature extraction methods, including the wavelet transform as discussed below.

Wavelet transform, due to excellent local zooming property of wavelet, is an effective tool for time-frequency decomposition for analyzing nonstationary signals. In this study, the multi-resolution wavelet analysis has been used to decompose the signal in time and frequency domain, while the continuous wavelet transforms of a continuous signal, $x(t)$, is defined by:

$$Wx(a,b) = x \otimes \psi_{b,a}(t) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t)\psi^{*}\left(\frac{t-b}{a}\right)dt \tag{1}$$

where ψ and ψ^{*} are the basic function and its complex conjugate; a and b are the scale and translation factors, respectively. Eqn. (1) is to decompose $x(t)$ into basic function $\Psi((t-b)/a)\Psi\left(\frac{t\text{-}b}{a}\right)$, named the mother wavelet. The scale factor a is equal to 2. The frequency spectrum of the wavelet is stretched by a factor of 2 and all frequency components shift up by a factor of 2. The discrete wavelet transform can be treated as a band-pass filter:

$$\mathrm{Wx(j,k)} = \int_{-\infty}^{+\infty} \mathrm{x(t)} 2^{\frac{j}{2}} \psi^{*}(2^{j}t - k)dt \tag{2}$$

Wavelet packet analysis behaves as a further generalized wavelet transform. It has different time-frequency windows to decompose signals, which are inconvenient in the wavelet decomposition. A wavelet packet function can be written as:

$$\psi_{j,k}^{i}(t) = 2^{\frac{j}{2}}\psi^{i}\left(2^{j}t - k\right) \quad i = 1, 2, \ \ldots, \tag{3}$$

where i, j, and k are the modulation, the scale, and the translation parameter, respectively. The ψ^{i} is obtained by using recursive relationship:

$$\psi^{2i}(t) = \sqrt{2}\sum_{k=-\infty}^{\infty} h(k)\psi^{i}(2t-k) \tag{4}$$

$$\psi^{2i+1}(t) = \sqrt{2}\sum_{k=-\infty}^{\infty} g(k)\psi^{i}(2t-k) \tag{5}$$

where $h(k)$ and $g(k)$ are the quadrature mirror filters. It is determined by mother wavelet function and scaling function. ψ^1 is the mother wavelet. The mother wavelet has some significant properties, including invariability and orthogonality. Wavelet packets have an adjustable time and frequency resolution. It has a different time and frequency resolution at every level. The top level has good resolution in the time domain and the bottom level has good resolution in the frequency domain. The frequency recursive relations are shown in **Fig. 2** for a full 3rd level wavelet packet decomposition, called the Mallat-tree decomposition.

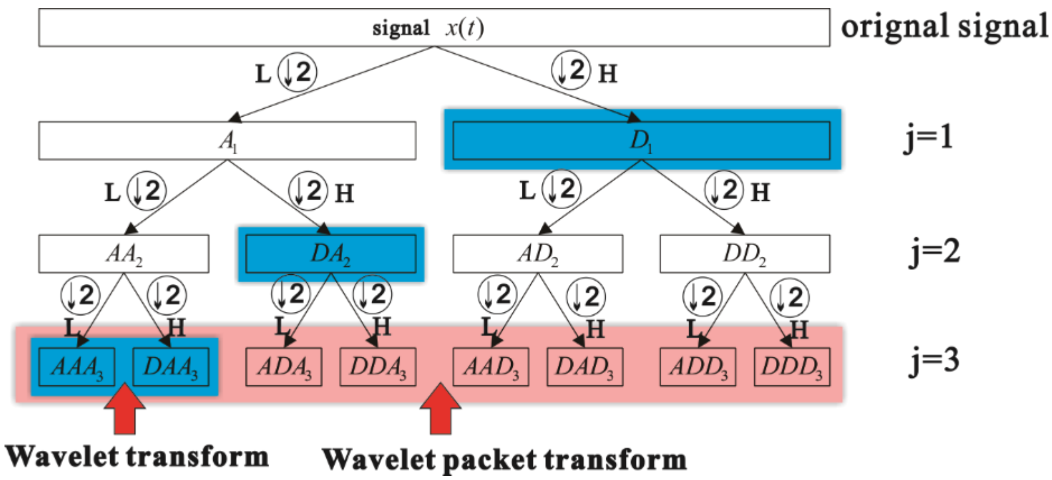


Fig. 2 3rd level wavelet transform and wavelet packet transform [3]

As illustrated in **Fig. 2**, the blue box and the pink box indicate the wavelet transform and wavelet packet transform of the signal, where *H* means high-pass filtering and *L* means low-pass filtering, *A* and *D* denote the approximation coefficients and detail coefficients, respectively. The recursive relations between the j^{th} and the $(j+1)^{th}$ level are by the form:

$$x_j^i(t) = x_{j+1}^{2i-1}(t) + x_{j+1}^{2i}(t) \tag{6a}$$

$$x_{j+1}^{2i-1}(t) = (x_j^i(t) * h) \ \downarrow 2 \tag{6b}$$

$$x_{j+1}^{2i}(t) = (x_j^i(t) * g) \ \downarrow 2 \tag{6c}$$

By using the inverse Fourier transform, Eqn. (6) is converted into the time domain as

$$Wx(a,t) = \mathcal{F}^{-1}\{Wx(a,f)\} \tag{6d}$$

where $\mathcal{F}^{-1}\{\cdot\}$ denotes the inverse Fourier transform. The variation of the scale factor, *a*, could yield different resolutions in different domains. A relatively small-scale factor could provide a high resolution in the time domain, while one could have the better resolution in the frequency domain with the increase of the scale factor. As a result, the continuous wavelet transform can generate the better adjustable time and frequency resolutions at any scale over other two methods. Note that the continuous wavelet transform will be later abbreviated as the wavelet transform for simplicity, unless otherwise noted.

2.4. Feature representation and classification using machine learning

2.4.1 Support vector machine (SVM)

SVM has been a powerful tool for classification problems in machine learning, which is developed by Vapnik [4]. The principle of SVM for classification is to construct a hyperplane which separates the data into two classes. It maps the input vector into a higher-dimensional feature space by applying kernel function (e.g. linear, polynomial or Gaussian radial basis function). Then, an optimal hyperplane is established in that feature space to make the separation which maximize the margin from the hyperplane to the closest data points in either class.

In general, three kernel functions tend to construct a higher dimensional feature space and allows a projectile of data to this hyperplane(s) to achieve being linearly separable. These kernel function helps SVM much more suitable for different dataset. The SVM can also be used in non-linear classification. The kernel function is introduced into SVM which could map the data points into a high dimensional space. The separating hyperplane is constructed in this space. There are some popular kernel functions, including linear function, polynomial function, Gaussian radial basis function and so on.

2.4.2 Optimization techniques

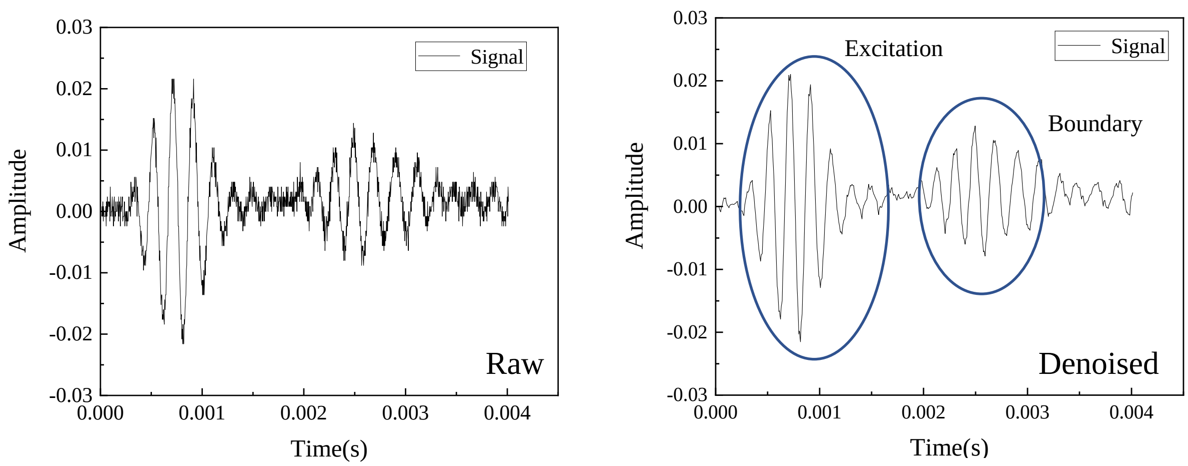
Selecting suitable penalty coefficient and kernel function parameter for the SVMs could enhance their accuracy for damage classification. Three representative optimization techniques are selected to optimize the parameters in the SVM: a) Grid-search techniques (GS); b) particle swarm optimization (PSO); and c) genetic algorithm (GA), which are addressed in detail below. Note that these three widely accepted approaches are selected for simplicity to demonstrate the proposed concept, although there are many optimization techniques in the literature, and they could be used to gain more information from optimization viewpoints.

3. Results and Discussion

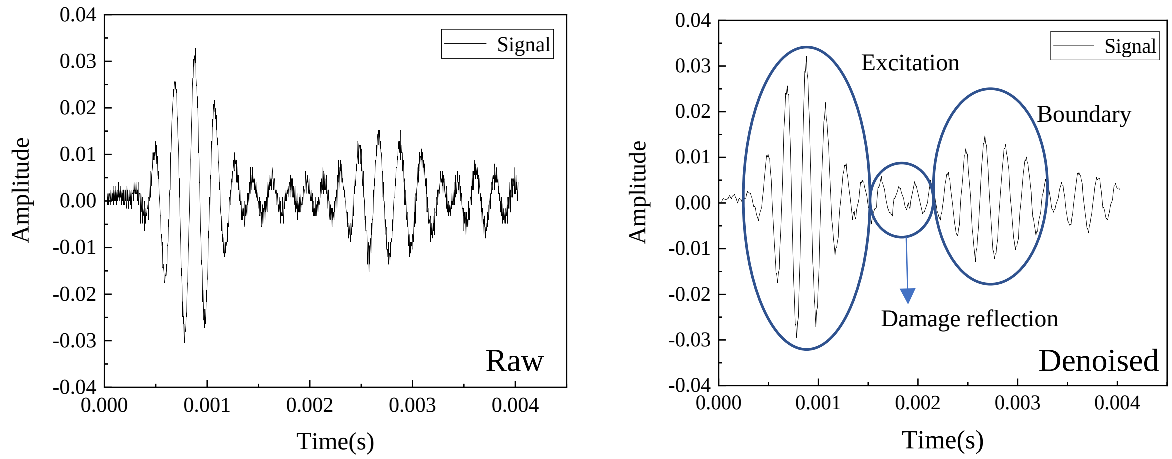
3.1 Limited experimental datasets

The lamb wave excitation was designed by a 5-cycle tone burst filtered through a Hanning window with 1 kHz to six steel plates under three different scenarios. Received raw signals were collected and de-noised by wavelet transform, as shown in **Fig. 3**. From time-domain standpoint, these signals exhibited similarity, but different trends. Signal under undamaged state displayed two main packages, representing initial disturbance and the reflection from the boundary. At the first package, the excitation signal was a 5-circle wave and its tails were the reflection from the boundary of the left side (which was near the actuators). The results revealed that the speed of the lamb wave was nearly 600 m/s.

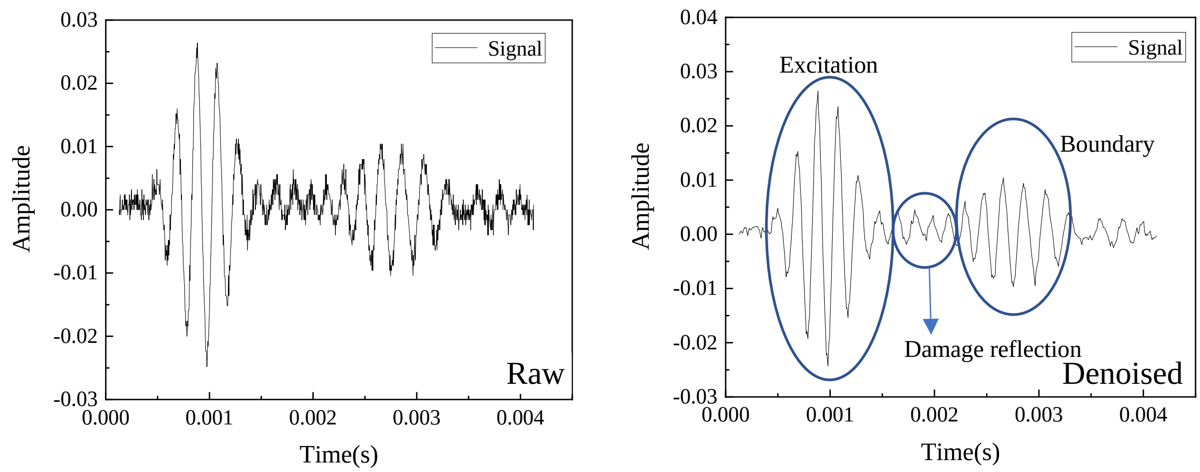
Differently, as shown in **Fig. 3(b)**, some reflections were observed between the first package and the second package, which was the response due to 20mm-long notch damage. Compared with **Fig. 3(d)**, the amplitude of the signal (b) that represented the damage area was slightly larger. When 20-mm long notch rotated 45 degrees, the reflection exhibited similar trend as the vertical one. The signal of the damage with circular shape has different waveform, as compared to the ones with notch shape at the response of boundary. **Fig. 3(e)** exhibited relatively bigger reflection at damage response as compared to **Fig. 3(f)**. Thus, the returned signals showed the difference distinctly due to different damage types, sizes, and orientations.



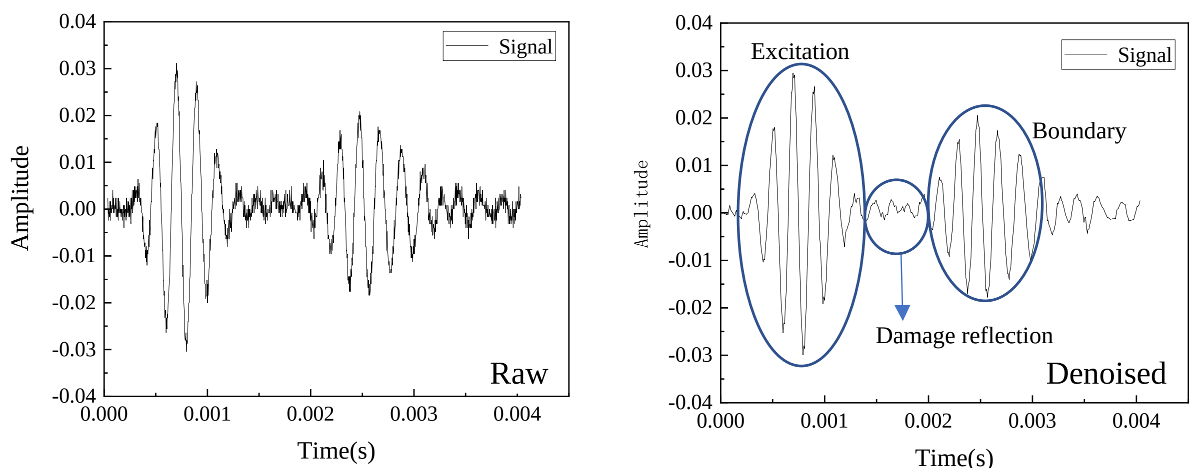
(a) Undamaged state



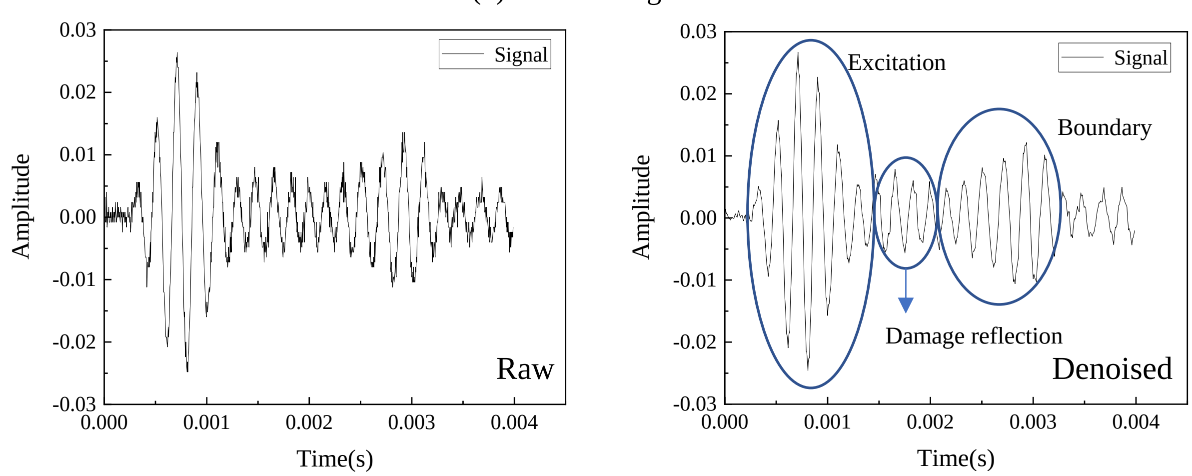
(b) 20 mm long notch state



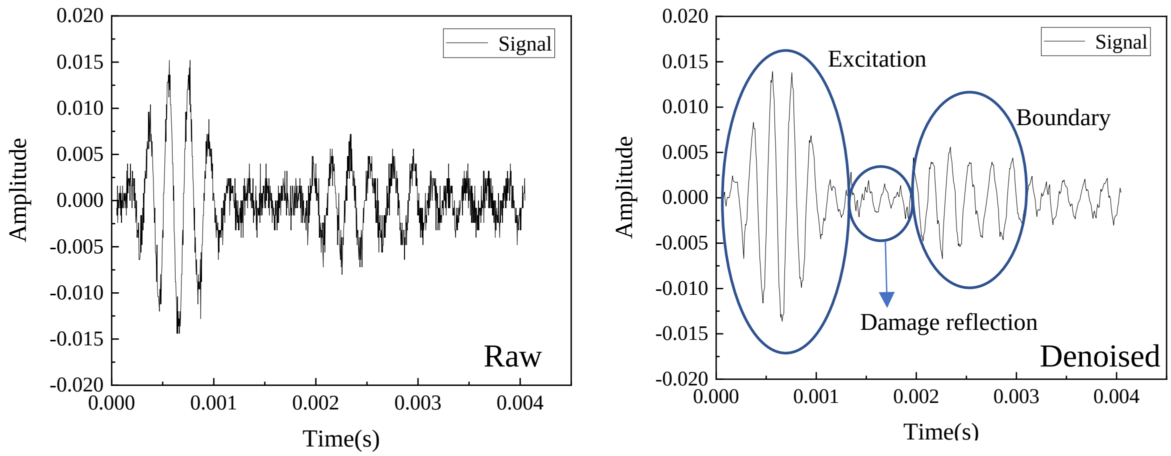
(c) 20 mm long notch with 45-degree rotation state



(d) 10 mm long notch state



(e) 20 mm diameter circular shaped damage state



(f) 10mm-diameter circular shaped damage state

Fig. 3 Received raw and denoised signals

3.2 Numerical simulated datasets

The signal with different damages were received respectively. To analyze the received signal clearly, the signals were cut off at the second echoed signal, shown in **Fig. 4**. The signal propagated in the plate and echoes when it arrived the boundary and crack. In order to reduce the complexity of the signal, only the first three wave packets received by the receiver was used to extract the features. The first signal packet presented the excitation. Then, the second wave packet represented the echoed wave from the damage. Moreover, the third wave packet showed the echoed signal from the boundary away from the receiver. The time span of the signal also provided the same result. The time that second wave packet appears was half of the time that the third wave packet appears, which was the same with that the damage was in the middle of the plate. Therefore, the second packet of the signal presented the information for the damage.

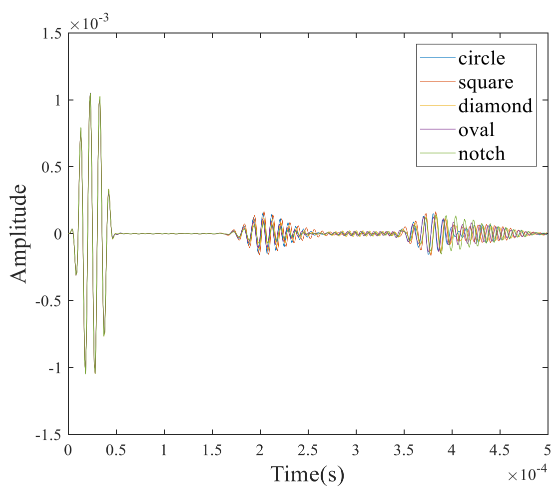


Fig. 4 Received signals through Pulse-echo method

In conclusion, data-driven approaches could help engineer to classify the damage. This method could consider the uncertainty (e.g., noise, measurement errors) that appear from environment and data collection. Moreover, noise interference could contaminate the data representation and in turn increase the risk of the data mining.

(c) Description of any Problems/Challenges

No problems are experienced during this report period

(d) Planned Activities for the Next Quarter

The planned activities for next quarter are listed below:

- First direction of the experimental tests will be conducted, while the 2D damage detection will be analyze with more damage types and uncertainty inclusion.
- Second direction of the algorithm of machine learning will be developed and 2D simulation will be set. Feature selection will be achieved by deep learning and parameter optimization.

Reference

[1] Gui, Guoqing, et al. "Data-driven support vector machine with optimization techniques for structural health monitoring and damage detection." *KSCE Journal of Civil Engineering* 21.2 (2017): 523-534.

[2] Ying, Y., Jr., J. H. G., Oppenheim, I. J., Soibelman, L., Harley, J. B., Shi, J., and Jin, Y. (2013). "Toward Data-Driven Structural Health Monitoring: Application of Machine Learning and Signal Processing to Damage Detection." Journal of Computing in Civil Engineering, vol. 27, pp. 667-680.

[3] Pan, H. , Azimi, M. , Yan, F. , and Zhibin Lin (2017) "Time-Frequency Based Data-Driven Structural Diagnosis and Damage Detection for Cable-Stayed Bridges." *ASCE Journal of Bridge Engineering*, In press.

[4] Valdimir, V., and N. Vapnik. "The nature of statistical learning theory. 1995."