Research Article | Vol. 27, No. 5 | 4 Mar 2019 | OPTICS EXPRESS 7685
Optics EXPRESS

Nuisance alarm reduction: using a correlation-based algorithm above differential signals in direct detected phase-OTDR systems

M. ADEEL,[1] C. SHANG,[2,3,*] K. ZHU,[1,3] AND C. LU[1,3]

[1] *Photonics Research Centre, Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China*
[2] *Photonics Research Centre, Department of Electrical Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China*
[3] *The Hong Kong Polytechnic University Shenzhen Research Institute, 518057, Shenzhen, China*
[*] *chao.shang@polyu.edu.hk*

Abstract: Significant research efforts have focused on techniques for alleviating the nuisance alarm rate (NAR) in the field of ϕ-OTDR pattern recognition systems. Unfortunately, ephemeral events were mostly neglected in previous research, and algorithms meant for improving classification accuracy were emphasized at the cost of acquiring a very large number of traces. This problem engendered an additional source of NAR in a specific class of events. The proposed solution uses a novel correlation based wrapper on top of differential signals that aims to filter out the effect of unnecessary phases in direct detected ϕ-OTDR systems. This technique avoids the use of irrelevant data in these differential signals by exploiting a better use of these unnecessary phases and provides a better intensity translation with fewer acquired traces as compared with contemporary techniques of extracting features.

1. Introduction

ϕ-OTDR uses phase change among Rayleigh Backscattered (RBS) signals as the only condition in measuring the disturbance along distributed fiber sensing system. Due to its low cost and relatively significant high sensitivity, it provides a range of applications including intrusion detection for security systems along a length of distance [1–4], warning systems [5], sensing for railway track systems [6], detecting damage tolerance of a system [7], and detecting ultrasonic waves [8]. It is believed that the ϕ-OTDR system has already attracted a great deal of study in the contemporary era. The effects like Rayleigh noise [9, 10] and non-coherent addition of RBS signal interference [11, 12] due to frequency drift of the laser source and laser linewidth are some of the sources of provision of random relation between input and fiber response which are responsible for a certain noise level in differential RBS signals. Even if the effects from the mentioned sources of nuisance alarm rate (NAR) are neglected, the vibration measurements are still non-comparable to those of conventional sensors. The reason is that the ϕ-OTDR response depends partly on an applied strain that is used to displace particles directly. Part of its dependence is upon the strain integration over the gauge length [13]. This part of response causes a large probability of differential signals to hide within the shadow of the noise and is regarded as a source of NAR for ephemeral events. Attempts were made to alleviate the effect of NAR by improving classification accuracy, but it led to an overall increase in computational cost [4, 14–16]. Further Steps were taken by Q. Sun [17] and X. Huang [1] for improving both classification accuracy and computational cost simultaneously at the expense of acquiring a myriad number of traces to detect a single activity.

The mentioned problem of phase integration along the gauge length becomes more dominant than the useful phase change due to the fiber stretch. This effect was possible by the consideration

#348713 https://doi.org/10.1364/OE.27.007685
Journal © 2019 Received 25 Oct 2018; revised 13 Dec 2018; accepted 13 Dec 2018; published 1 Mar 2019

Optics EXPRESS

of a very few number of traces as this small percentage of useful data within a small number of acquired traces is more certainly insufficient to tell different activities apart. This problem enhanced NAR in continuous type of events, events with no idle data in between successive impacts, for which an event lasts over a very short duration of time, whereas, most of the contemporary feature extractors in ϕ-OTDR pattern recognition systems were designed for an extremely large number of acquired traces. These conditions of uncontrollably less number of traces are more frequently raised in the non-continuous type of perturbations where a single impact of the applied force is insufficient to create a significant amount of traces with useful information, called active region. A simple reason for not curtailing the huge number of traces in contemporary techniques was the achievement of a significant classification accuracy, the ramifications of which were the involvement of idle traces along-with the active traces. Acquiring too many traces for recognizing a single event not only increases NAR by considering non-useful idle traces but this practice restricts the system to learn only long-lasting events along with an increase of the requirements of number of traces. For instance, considering 20,000 traces in [4, 16, 18] means an overall delay of 20sec other than processing delay stage if a 100km fiber is used under test. Due to huge data manipulation in real time in ϕ-OTDR systems, the use of feature extractors like Short-Time Fourier Transform (STFT) and Level Crossing (LC) were quite common due to their ease of simplicity [19] even after the introduction of complex feature extractors like EMD [16] or morphologic [17] based extractors. According to the latest review [19], the simple feature extractors like LC and STFT have been used about 82% of the time of all the feature extractors used. Unfortunately, almost a high percentage of all these feature extractors are based upon differential signals. The mentioned possibility in an increase in NAR was due to the direct involvement of these differential signals.

In case of non-continuous events like walking, running or jumping etc, the frequency of impact (f_p) and duration of impact (τ_p) are the two important parameters that are most affected by the contemporary feature extractors as these extractors are based directly upon differential signals. An attempt to improve f_p was made by Mahmoud in [4] but the use of LC feature extractor led to extreme noise levels and hence eventually enhanced the effect of NAR. With the use of feature extractors directly on the top of differential signals, the parameters f_p and τ_p were impossibly estimated with these contemporary practices due to the mentioned possibility of NAR. As we know that the boundaries of an active or useful region of the acquired data can be successfully constructed with the help of exact estimation of parameters f_p and τ_p. Pattern recognition is then applied to such a useful region to avoid the use of idle data which is normally existed between applied impacts. Normally, the contemporary used feature extractors were based on estimating the classification accuracy without estimating the mentioned parameters which enabled these systems to consider idle data and hence a decrease in requirement of total number of traces was possible.

These issues were resolved with the help of our proposed solution that relies upon the use of adjacent correlation operator, termed as a wrapper in this paper, that runs on top of the differential signals. This practice filters out the effect of integration of phases over the gauge length and leaves behind the effect of phase change due to the fiber stretch only. With this technique, there is a possibility of exact estimation of the parameters f_p and τ_p and a possibility of recognizing the pattern of only active traces without involving the use of irrelevant data. This solution can be considered as a significant improvement in the applications involving pattern recognition without imposing restrictions on ϕ-OTDR design. The proposed solution avoid the extraneous use of acquired traces in two ways. First, it extracts useful data from a very short length of acquired traces to help at both parameter estimation and pattern recognition stages and secondly it can possibly avoid the use of non-idle data at the pattern recognition stage. In future, the use of proposed wrapper on top of differential signals is expected to be used for recognizing different events based on patterns within active region only in assistance to the parameter estimation of

f_p and τ_p which is expected to mitigate total delay tremendously along with providing better intensity translation of the applied perturbations.

In short, despite of practicing the traditional way of increasing classification accuracy, the emphasis has been provided to enhance this accuracy by eliminating NAR in all possible event types effectively. The possibility of NAR was due to the involvement of small number of traces and intrinsic non-linearity between input and output response which in-turn affected both parameter estimation and pattern recognition. Improvement of parameter estimation as an application example has been explained in this paper. The additional advantages of using the proposed feature extractor include a more probable possibility of improving pattern recognition in the active region only, the use of a direct detected ϕ-OTDR system on the account of widely installed similar systems and a possibility of use of contemporary algorithms based on these systems [2, 15–19].

2. Experimental setup and methodology

A simple direct-detected ϕ-OTDR system configuration has been implemented with a Laser source of 1550nm wavelength with less than 0.1KHz linewidth. A 7000m long Fiber Under Test (FUT) was used under different experiments using PZT and real-life experiments. The Pulse Repetition Rate of 10KHz was considered in all the experimental trials, as shown in Fig. 1.

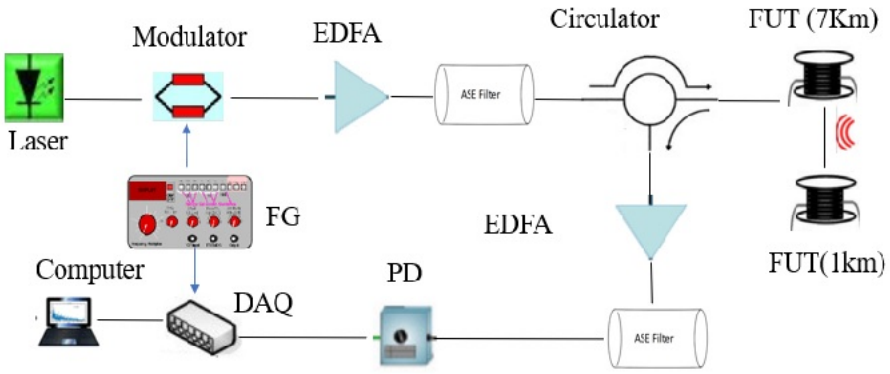


Fig. 1. Experimental Setup: Direct detected Phase-OTDR system.

To demonstrate the importance of our proposed feature extractor, let us consider the electric field of the RBS signal as a result from a traditionally used rectangular pulses injected at z=0 which is shown in Eq. (1).

$$E(z) = E_0 e^{(-2\alpha z)} e^{(j\omega t_0)} \sum_{u=1}^{U} r_u e^{j\phi_u} \tag{1}$$

where α is the attenuation constant that cause the signal to degrade along the fiber of length z, and ϕ is the phase relative to each RBS signal. The electric field intensity of the i^{th} trace can be defined as,

$$\gamma_i(n) = \mid E \mid^2 = \sum_{u=1}^{U} r_m^2 + 2 \sum_{\upsilon=1}^{U-1} \sum_{u=\upsilon+1}^{U} r_u r_\upsilon cos(\phi_u - \phi_\upsilon) \tag{2}$$

The differential signal is obtained after subtracting each subsequent trace from the previous one as shown,

$$\delta_i(n) = \gamma_i(n) - \gamma_{i+1}(n) \tag{3}$$

Optics EXPRESS

where i represents the trace number which goes from 1 to the length β and n shows a specific data-vector within the resolution cell of the perturbed region and it ranges from 1 to N_r. It implies the differential equation between any two traces and under a single the umbrella of a pulse without considering its drag along the fiber as,

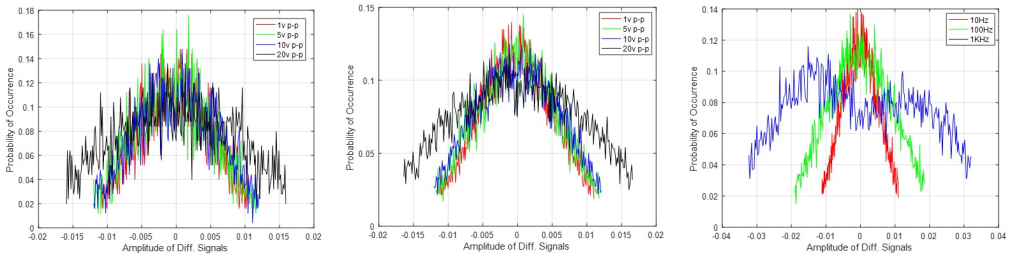
$$\delta = 2\sum_{\upsilon=1}^{q-1}\sum_{u=q}^{U} r_\upsilon r_u \left[cos(\phi_\upsilon - \phi_u) - cos(\phi_\upsilon - \phi_u - \varphi_p)\right] \tag{4}$$

where the angles ϕ_υ and ϕ_u represents a set of angles from each contributing regions of perturbation and the integration of these angles give rise to a response that affects the phase φ_p. The angle φ_p is engendered due to the axial displacement of particles within the fiber.

As the pulse traverses the perturbed region, the effect of summation of the BRS signals across the resolution cell varies against each n in Eq. (4) but its response to a change in fiber length due to external perturbation is almost constant for each n except the regions with strong coherence fading. The combined effect of both these type of phases has resulted in a large probability of non-linear response of the differential signals. This non-linearity normally propagate after the transformation of data is made with the use of traditional feature extractors. For instance, let us assume the differential signals from Dataset 1, Dataset 2, Dataset 3 and Dataset 4 of [20], with 5,000 and 20,000 traces as shown in Figs. 2(a) and 2(b) respectively. It can be observed from these figures that the response of a very small probability of differential signals get increased with an increase in applied intensity. This discrimination is more prominent as the number of acquired traces or the intensity of the applied perturbation is increased. The higher the intensity of the applied perturbation, more dominant will be the response of differential signal due to fiber stretch as a large percentage of raw signals exhibit instant jumps due to the inherent periodic nature of the ϕ-OTDR system [13]. This can be thought of a noise threshold against differential signals above which one can see a very less number of a dominant signal for each increase in applied intensity. From the perspective of a common observation, a very high probability of differential signals share the common levels of amplitude, the occurrence of which is closed to zero. A very less probability of these signals represents the level with which the applied 4 different perturbations can be differentiated. The LC feature extraction algorithms based on this rule require too many traces to get a significant amount of useful levels in order to differentiate between these activities. Let us consider another case in which a peak-peak voltage of the applied square signal is 20v and the frequency of these signals are varied with the help of PZT as shown in Fig. 2(c) with the data of interest available at Dataset 5, Dataset 6, and Dataset 7 of Dataset 1, Ref. [20]. We can observe that the three distinct differential response signals can be discriminated with respect to the center frequency. The level of these response signals gets increased with an increase in center frequency as for large frequency signals the differential signals cross the level above the noise threshold more frequently, resulting in a large number of signals above this threshold. In other words, the signal traverses the threshold more frequently, beyond which the phase change due to the fiber stretch become more dominant than the response due to phase integration along gauge length. But again, the probability of the number of levels used for discriminating these signals is the way much less than the probability of their common amplitude levels.

In order to get a very high discrimination power among different signals, the emphasis was given to acquiring a very large number of traces for getting a reasonable amount of useful data. Useful data means the differential data whose levels do not overlap at the output against different perturbations. This data was then used as a source of discrimination among different intensity or frequency perturbations. There was a need to transform the differential signals in a way which could only represent the response to the fiber stretch only. To eliminate the effect of non-useful phases which are responsible for non-faithful response, a correlation operator is applied among each data-vector within the resolution cell of the perturbed region. This operator acts like a matched filter which normally mitigates the effect of intrinsic phases, leaving behind the effect

Optics EXPRESS



(a) (b) (c)

Fig. 2. Probability of occurrence of each differential signal amplitude level out of total 256 levels after applying the simulated perturbation using, (a) Sinusoidal signal with 100Hz frequency and considering 5000 traces, (b) Sinusoidal signal with 100Hz frequency and 20,000 traces, (c) Square signal with 20v p-p voltage to PZT and considering 20,000 traces.

due to the phase change affected by the applied perturbation only. This effect can be thought of a wrapper on the top of differential signals which can be served as a useful feature extractor or a base signal for any feature extractor.

To elaborate on the concept of proposed feature extractor or the wrapper in a more convenient way, let us consider the number of traces to represent a vertical distance and the spatial length to show the horizontal distance. Each data-vector is defined to represent the β number of traces. A single data-set represents the results of computations made by each sliding window with vertical length β and the horizontal length N_r drags vertically all the way through the entire data-set of length T_t within the perturbed resolution cell. The sliding window takes an incremental jump α for each move during sliding. Before any subsequent jump by the dragging window, the system computes the Pearson correlation coefficient as applied to the adjacent data-vectors δ_i, which can be expressed as,

$$R_j[n] = \frac{\beta \sum_{i=\alpha_j}^{\alpha_j+\beta} \delta_i(n)\delta_i(n+1) - \left[\sum_{i=\alpha_j}^{\alpha_j+\beta} \delta_i(n)\right]\left[\sum_{i=\alpha_j}^{\alpha_j+\beta} \delta_i(n+1)\right]}{\{\beta \sum_{i=\alpha_j}^{\alpha_j+\beta} \delta_i^2(n) - \left[\sum_{i=\alpha_j}^{\alpha_j+\beta} \delta_i(n)\right]^2\}^{1/2} \times \{\beta \sum_{i=\alpha_j}^{\alpha_j+\beta} \delta_i^2(n+1) - \left[\sum_{i=\alpha_j}^{\alpha_j+\beta} \delta_i(n+1)\right]^2\}^{1/2}} \tag{5}$$

$\forall j = 1, 2, 3, \cdots J$. It implies $\alpha_J + \beta < T_t$, where T_t is considered to be the total number of traces in a single data set. For a significant applied intensity, every adjacent data-vector in the perturbed resolution cell is different by the magnitude of coherence fading only. The overall trend of every data-vector is changed proportionally due to a strong relative relation among these vectors. Due to these reasons, the simple correlation operator like Pearson correlation was suggested.

After the processing window traverses the whole data-set, it computes a set of vectors represented by $\mathbf{R}_j$ with each j^{th} vector is comprised of $N_r - 1$ elements. A matrix $\mathbb{R}$ can be defined to contain the whole set of vectors $\mathbf{R}_j$ such that $\mathbb{R} = \mathbf{R}_j, \forall j$. The collective process of exploiting multiple domains and taking adjacent correlation for each subsequent data-vector within the resolution cell is termed as the proposed feature extractor or a useful wrapper on the top of differential signals for further processing.

The purpose of computing the vector $\mathbf{R}_j$ against a specific j^{th} window eliminates the effect of an ambiguous response of the ϕ-OTDR system due to intrinsic phases and the matched filtering provides a proportional response between input and the transformed data. Part of the signal that is changed due to fiber stretch is common among all data-vectors within the resolution cell and this matched filter provides a higher value of correlation for the displacement of the particles within fiber due to axial stretch. Part of the signal that is engendered due to an integration of phases among gauge length is different for each data-vector. This claim can also be justified after

Optics EXPRESS

the correlation operator in Eq. (5) is applied among all N_r data-vectors against a specific j^{th} movement of the sliding window. Figure 3 shows an increase in correlation among data-vectors within the resolution cell of spatial points between 13 and 50 against a spatial resolution and the gauge length each with the length of 10m. Each spatial point represents the resolution of a DAQ card with a speed of 250MS/s. As we see the gauge length equals that of spatial resolution, yet the response is quite proportional. The figure demonstrates quite a proportional amplitude response of the transformed data after the proposed feature extractor is applied for each increase in peak-to-peak voltage by the applied simulated perturbation for a very small number of traces.

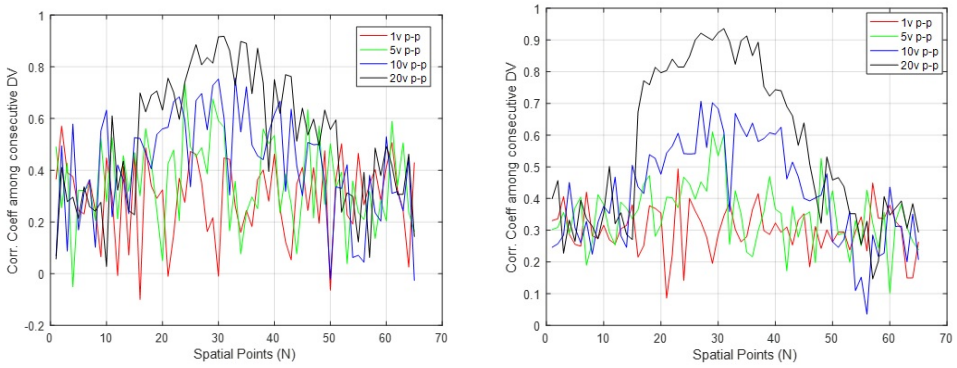


Fig. 3. Demonstrating response of the proposed wrapping algorithm for 4 different peak-peak sinusoidal signals using, (a) β=50 (b) β=200.

Figure 3 shows a very clear demonstration of discrimination among different activities based on their intensities for a very low number of traces with the help of running the wrapper which is based on the correlation among individual data-vector. Figure. 4 considers all the possible values within the active region of Fig. 2 and plots the probability density of each of these values with a precision of up-to two decimal points. The probability of non-useful differential traces is apparently larger than the probability of useful traces, as in Fig. 2, the mitigation of dependence on intrinsic phases by the proposed wrapper provides the same probability difference in the opposite manner. In other words with this solution, the probability of useful differential traces

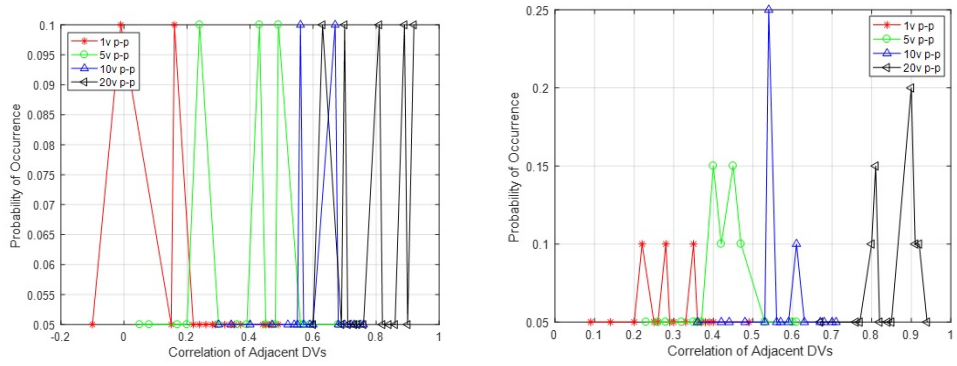


Fig. 4. Probability of occurrence of the correlation among data-vectors for, (a) β=50 and (b) β=200.

Optics EXPRESS

is larger than the probability of non-useful differential traces for very less number of traces, as shown in Fig. 4. It shows a majority part of distinction among different perturbations does not overlap after the wrapper is applied on the top of differential signals. The overlap can only be found in the low probability density parts as the major dependency of the response is due to particle displacement in FUT. This figure also represents insignificant change in distinguishing amplitude levels if β is increased from 50 to 200, showing a great potential of our proposed solution at the very less number of traces. The figure unveils this distinguishing characteristic of the proposed solution for the DAQ resolution of 20 points within a spatial resolution cell of 100ns with a 10m fiber length exposed under perturbation.

3. Implementation example: parameter estimation

A fascinating advantage of the proposed feature extractor about eliminating the effect of integration of phases over a gauge length has been explained in the previous section. Other advantages of the same characteristic are applicable to two types of activities. The first type includes a true intensity translation in parameter estimation. This estimation can be used in non-continuous perturbations only where the active region of perturbation is separated from the idle state of the acquired traces by extracting the parameters f_p and τ_p. The second application where this proposed feature extractor can be used is feature estimation or the pattern recognition by enabling the system to use a fewer number of traces to achieve the same classification accuracy as achieved by the traditional feature extractors. This can be done by exploiting the active region of perturbations only in the non-continuous type of activities or by using the whole series of traces in ephemeral events with the continuous type of activities. A solution to improve the NAR from a number of distinct perturbations of non-continuous nature is discussed in this section by introducing the concept of prior probabilities which are based on two important parameters, termed f_p and τ_p. This can be thought of an application example of our proposed wrapper.

The events with dynamic perturbation applied on FUT with an idle gap of no activity between the applied impacts are termed as non-continuous, whereas, for no idle gap of activity the said activities can be classified as continuous. The parameters f_p and τ_p can only be estimated in the non-continuous type of events. Predicting the mentioned parameters of a certain activity accurately aids in detecting active region of the data and a pattern recognition scheme can only be applied to the fetched data after the active region is detected. By using data from the active region only, there is no need to feed irrelevant or idle data between successive impacts of the activity. In other words, the parameter estimation can be regarded as a first step in a hierarchy to classify the events on broader sense before further classification is done with the help of pattern recognition. It is a very effective way of increasing classification accuracy without relying on a myriad number of traces. Hence, a very small number of traces can be used as an alternative to a very large number of traces with the help of proposed solution by first increasing the probability of useful traces and then avoiding the use of non-idle data at the pattern recognition stage.

From the previous discussions, we can see a single correlator was used as the feature extraction operator which clearly indicates its lightweight nature. This solution provides a very high freedom of selection of the quantities α and β without affecting parameter f_p even for the high frequency of impact. If β is exceeded such that its length covers two impacts within a single sliding-window, it leads to the inaccuracies. Hence, it is always appropriate to keep this value shorter based on the highest value of the parameter f_p. The factor α represents the jump each sliding window undertakes during its entire motion within a single block of a specific data-set. Keeping this value to its lowest improves the f_p as well as τ_p for a certain activity. However, a very small value of α should be a trade-off as it increases the computational cost of the algorithm.

To calculate the best estimates for both f_p and τ_p, selecting a right quantity of α and β requires complex iterative methods regarding changing their values according to each type of external activity. Keeping appropriate values for both α and β is challenging as the duration of each q^{th}

Optics EXPRESS

impact, termed as $\tau_p(q)$, varies for a certain type of p^{th} activity. For instant if we assume the traditionally used feature extractors to neglect idle data and make computations based on the data of active region only, it may thus require a very careful selection of the term β as these extractors are normally required to do computations based on a single dimensional data which is considered to be quite limited amount of data for computing fruitful results.

To consolidate our claims let us first consider the data from two different activities. These activities were classified as slow f_p and high f_p and were meant to change their location for each impact within the range equal to spatial resolution and can be considered as analogous to walking and running respectively. For the sake of comparison with the proposed feature extractor, we have applied both LC and STFT to clearly demonstrate the benefits associated with this solution.

Let us consider LC with the Number of levels (N_l). The variable $L_j(p)$ represents the number of levels at j^{th} window and p^{th} amplitude level as given.

$$L_j\left[p\right] = \sum_{t=\alpha_j}^{\alpha_j+\beta} f\left[\delta(t) \geq \varepsilon(p) \;\; \& \;\; \delta(t-1) < \varepsilon(p)\right] \tag{6}$$

where the function $f(y)$ is either 1 or 0 for true and false conditions respectively, $\forall t$ goes from 0 to $\beta - 1$ and p counts from 0 to $N_l - 1$, which is a user-defined quantity. The term $\varepsilon(p)$ represents each p^{th} threshold.

By taking NFFT points in-case of STFT which is assumed to have N_f points for each j^{th} computing window, the relation is given as.

$$F_j\left[\omega\right] = \sum_{t=\alpha_j}^{\alpha_j+\beta} \delta(t) e^{-j\omega t} \tag{7}$$

where ω goes from 0 to $N_f - 1$. $\forall t$ goes from 0 to $\beta - 1$ for a specific data vector $\delta(t)$.

To apply maximum likelihood estimation (MLE) based classifier, the feature extractor must have a very high noise resistance, i.e, there should be a high freedom of selecting both the useful parameters like α and β in order to keep the noise level at its lowest. In other words, selecting a simple threshold quantity of these parameters is required to better identify the perturbed region within the data if either of these feature extractors (LC and STFT) have the low noise resistance characteristic. To calculate MLE, the average of N_l levels are taken under each move of the sliding window in case of LC. For STFT, the average of the computed vector under each sliding window movement, equal to N_f points, is considered. In this case, only those indices are considered for which the FFT value is greater than a specified threshold. Before applying MLE based classifier for extracting f_p and τ_p, we consider the parameters θ_{NP} and σ_{NP} to be the average and standard-deviation respectively for data computed at the end of each step of sliding-window in the non-perturbed region. Considering the same concept of dragging window and computing the average θ_{ML} of the sliding-window which is dragged at its j^{th} location, can be defined as,

$$\theta_{ML}(j) = arg(\max_{\theta_{NP}\,\epsilon\,\boldsymbol{\theta}} p(\theta(j))|\theta_{NP}, \sigma_{NP}) \tag{8}$$

The term $p(\theta(j))|\theta_{NP}, \sigma_{NP})$ represents the likelihood of a certain perturbation against both perturbed and non-perturbed regions. It is observed that the vector sets $\mathbf{L}_j$ and $\mathbf{R}_j$ of the non-perturbed region follow the Gaussian distribution against each vector j as shown,

$$p(\theta(j)|\theta_{NP}\,,\;\sigma_{NP}) = \frac{1}{(2\pi\sigma_{NP}^2)^{J/2}} exp(-\frac{1}{2\sigma_{NP}^2}\sum_{j=1}^{J}[\theta(j)-\theta_{NP}]^2) \tag{9}$$

Optics EXPRESS

where, the mean θ_{NP} and the standard deviation σ_{NP} can be mathematically expressed as,

$$\theta_{NP} \approx \frac{1}{J}\sum_{j=1}^{J}\theta(j) \tag{10}$$

$$\sigma_{NP} \approx \sqrt{\frac{\sum_{j=1}^{J}[\theta(j)-\theta_{NP}]^2}{J}} \tag{11}$$

For the perturbed resolution cell, $\theta_{ML}(j) \neq \theta_{NP}, \forall j$, whereas, for the non-perturbed resolution cell, $\theta_{ML}(j) = \theta_{NP}$, assuming β to be equal to the length represented by $\tau_p(q)$, the parameter f_p can be calculated as the repetition rate of the occurrence of $\theta(j)$ for $\theta_{ML}(j) \neq \theta_{NP}$. The term τ_p for a certain perturbed cell can be measured for the same $\theta_{ML}(j)$ and this duration can be expressed as the interval between the boundaries of the index j of $\theta(j)$ for $\theta_{ML}(j) \neq \theta_{NP}$.

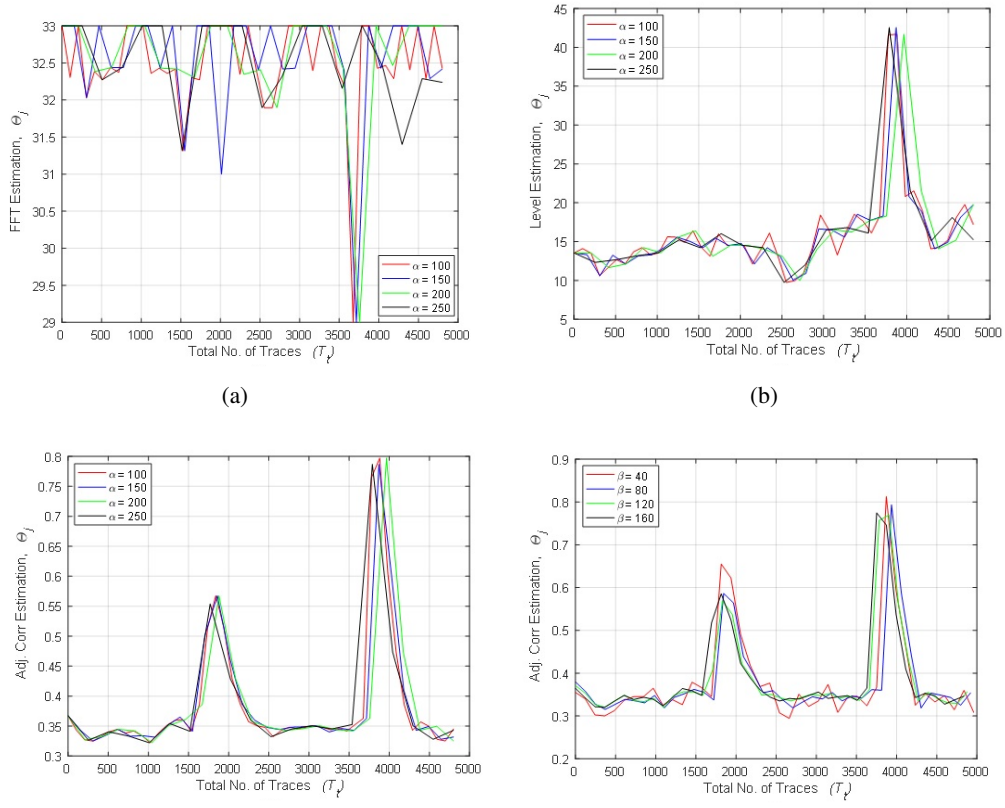


Fig. 5. Estimation results applied as a result of Walking activity, If (a) STFT FE is employed such that α is varied, keeping β=200, (b) LC FE is employed such that α is varied, keeping β=200, (c) Proposed wrapper is employed such that α is varied, keeping β=200, (d) Proposed wrapper is employed such that β is varied, keeping α=120

In order to demonstrate the noise resistance of the traditionally used feature extractors, termed LC and STFT, we first consider these algorithms to represent each of the walking and running activities. The number of active traces under each impact is as low as around 100 and this number can be further decreased if the pulse repetition rate is lowered due to an increase in FUT length.

Optics EXPRESS

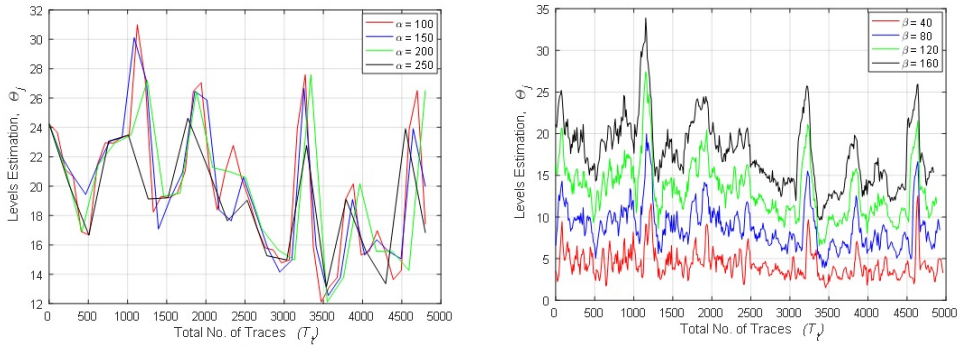
Assuming the single step process of computing LC and STFT to be the same, the results are demonstrated for a range of values from both the parameters α and β. Let us consider the activity of low f_p analogous to that of Walking. The frequency of impact, in this case, is restricted to 4Hz and the location of the perturbation is considered constant as the whole activity has been accomplished within a length much lesser than the spatial resolution of 48m. Considering 5000 traces against two impacts of Walking activity, as in Dataset 9 [20], around the location of traces that traverses 1700 and 3700 in number. Each former impact is less intensified than its subsequent counterpart. Figure 5(a) shows the estimate of a computational result of FFT from each j^{th} window after STFT feature extractor is applied. Part (b) of the same figure shows the estimation results of the levels attained after LC feature extractor is applied. Benefits of STFT over LC feature extractor is that each former step of the applied impact force is visible which was engulfed within the noise level in case of LC feature extractor. One of the drawbacks of using STFT feature extractor is that the location of perturbation is diverted from its true value against a certain value of α which shows its low freedom of selection of α. Figures 5(c) and 5(d) demonstrate the adjacent correlation estimated against each sliding window if the proposed feature extractor is implemented. The figure clearly depicts the position and amplitude of each former low intensified impact. A high freedom of selection of both α and β in case of proposed feature extractor provides an edge to its wide scope of adoption in future applications where boundaries of the active region of each applied impact are solicited.

In a similar way, the seven different impacts with f_p=14Hz, analogous to Running activity were demonstrated against each mentioned feature extractor. Unlike low f_p, the results get much worse if f_p is increased. As STFT exhibits very small freedom of selection of α and β, a small diversion of the impacted region has resulted in a more scrambled representation of the activity. LC feature extractor does not exhibit any high diversion which leads to its possibility of representing the activity with high f_p to a significant level of accuracy. Figure 6 demonstrates some ambiguity in these results after LC feature extractor is employed by considering the data Dataset 8 in [20]. This ambiguity is due to the fact that the response of particle displacement in the ϕ-OTDR system is dependent partly on the integration of strain over the gauge length. LC feature extractor takes this dependent factor into account and hence the response is not faithful to the intensity of applied perturbation. Submergence of low-intensity perturbation within noise level in Fig. 5(b) was due to the same reason. Because of such a response, a single threshold line cannot be attributed to differentiate between active and non-active regions of the activity using Running activity. Some applications might require a diverse range of values of β due to a variation in $\tau_p(q)$ where it is very hard to differentiate between active and non-active regions in case of LC as the level of each extracted data gets changed with a change in the parameter β. A simple algorithm like MLE is thus impossible to be applied. The implications are that the complicated algorithms might be used to tackle this problem but in such a case certain restrictions might be applied to restrict the application to few types of activities. A freedom to select the appropriate threshold is thus affected in case of applying LC extractor.

Figure 7 shows the application of the same data but with proposed feature extractor. The average of each sliding window shows a smooth representation of each applied impact in this case. Increasing f_p does not affect the overall extraction for the same range of values of α and β as were applied in case of STFT and LC feature extractors. After applying the MLE algorithm on proposed feature extractor, it is observed that a simple threshold of 0.6 is enough to discriminate between active and non-active regions for a range of quantities α and β.

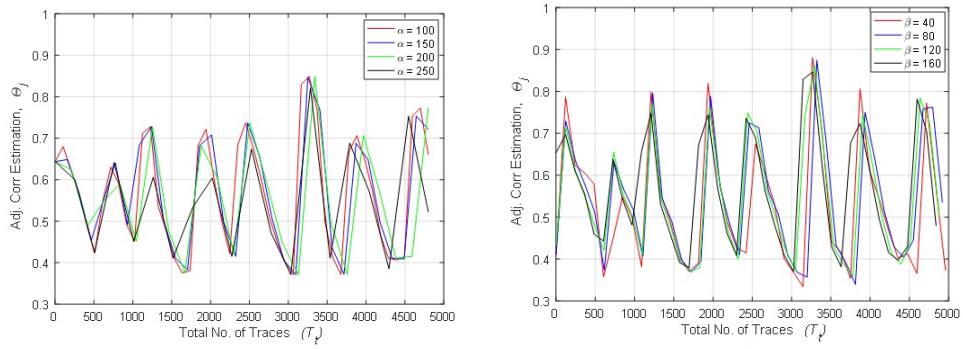
Table 1 shows a very slight difference in the resultant error for each q^{th} impact in every derived quantity of $\tau_p(q)$ for a free selection of a range of values of α. Keeping a simple threshold of $\theta_{NP} < 0.6$, no error in f_p was observed for a range of α from 30 to 150 by using $\beta = 150$. The last row of the same table clearly indicates a decrease in overall processing time for an increase in α. Unlike the contemporary feature extractors, a very high freedom of selection of α

Optics EXPRESS



(a) (b)

Fig. 6. Estimation results applied after LC feature extractor is used, If (a) α is changed, keeping β=200, (b) β is changed, keeping α=120.



(a) (b)

Fig. 7. Estimation results applied after the proposed feature extractor is used, If (a) α is changed, keeping β=200, (b) β is changed, keeping α=120.

reflects absolutely no change in f_p with a very slight difference in the values of τ_p, and this large freedom of selection is known to save much of processing time with the possibility of selecting a reasonable quantity of the term α.

The experiments demonstrated in this section gives a very clear idea about preferring proposed wrapper or feature extractor over traditionally used feature extractors. This preference is due to its higher noise resistance characteristics and due to better SNR, the false negatives can easily be countered with our solution. Second, a high freedom of selection of the parameters α and β was possible with our proposed solution and hence a very simple MLE based classifier is sufficient to provide a record level of accurate parameters of interest, i.e, f_p and τ_p. This high freedom of selection has removed the problem of existence of extreme levels of false positives for input perturbations with a high frequency of impact as it was mentioned in the example of activities analogous to running. A decrease in both false negatives and false positives by our proposed solution can decline NAR to an extreme level during the parameter estimation stage. For quantitative analysis of reduction in NAR, we can see the processing delay by our proposed solution can be reduced from 20.11 seconds to 4.74 seconds, as in Table 1, with a very minute error parameter estimation. which can also be observed from the results of the Fig. 7. Figure. 6 provides endless number of possibilities of NAR for a range of threshold levels in MLE estimation. For this reason, a random forest classifier was used ahead of the results of the STFT

Optics EXPRESS

Table 1. Error in seconds for each step for $\theta_{NP} < 0.6$ and $\beta = 150$, Running example with seven impacts in each trial if used with the Proposed Feature Extractor

	$\alpha = 30$	$\alpha = 60$	$\alpha = 90$	$\alpha = 120$	$\alpha = 150$
$\epsilon\left[\tau_p(1)\right]$	0.004	0.004	0.004	0.004	0.005
$\epsilon\left[\tau_p(2)\right]$	0.0048	0.0048	0.0108	0.0108	0.0048
$\epsilon\left[\tau_p(3)\right]$	0.0036	0.0036	0.0036	0.0096	0.0096
$\epsilon\left[\tau_p(4)\right]$	0.0025	0.0025	0.0005	0.0085	0.0025
$\epsilon\left[\tau_p(5)\right]$	0.0001	0.0031	0.0001	0.0031	0.0001
$\epsilon\left[\tau_p(6)\right]$	0.0046	0.0046	0.0046	0.0046	0.0014
$\epsilon\left[\tau_p(7)\right]$	0.0031	0.0001	0.0029	0.0001	0.0029
Proc. Time	20.11	10.32	7.27	5.87	4.74

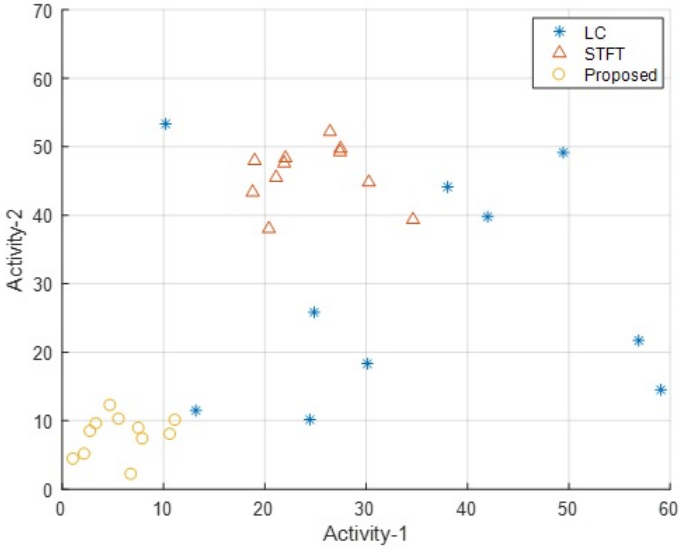


Fig. 8. Each axis represents error in percentage as False Positives by two non-continuous activities as (i) Act-1: Running and (ii) Act-2: Walking

and LC based feature extractors to better demonstrate the comparison. Fig. 8 provides the false positives associated with non-continuous events and the false positives in this figure provides a better quantitative comparison of our proposed solution as compared to the contemporary techniques of applying feature extractors on top of differential signals.

Optics EXPRESS

4. Conclusions

A large percentage of perturbed differential signals cannot contribute at any pattern recognition system due to a very small fraction of useful data in these signals. Despite increasing the number of acquired traces to get a reasonable number of useful data, a wrapper was designed on top of the differential signals to make use of existing low number of traces in a useful way of representing a certain perturbation. The correlation operated between every two adjacent data-vectors within the resolution cell of perturbed region was proved to directly translate the intensity of perturbation in a useful amplitude by filtering out the unnecessary output response in any ϕ-OTDR system. During the parameter estimation stage, the proposed wrapper was proved helpful to faithfully represent the difference between the intensity of each applied impact on the fiber with the help of a very less number of acquired traces. Due to its exploitation of additional domain, the wrapper aided in increasing SNR at parameter estimation stage and hence mitigated the number of false negatives. The number of false positives were proved to have been significantly decreased as compared to the traditional practice of utilizing the feature extractors. Overall, the NAR was dropped significantly with a decrease of both false positives and false negatives in a simple setup of ϕ-OTDR system. The same characteristic of the proposed solution can be proved helpful in slumping off NAR at pattern recognition stage in further research based on the proposed wrapper which will also decrease the need of idle traces. This wrapper can also be used directly as feature extractor or may be served as a basis algorithm for other feature extractors while considering data from the active region at pattern recognition stage.

Funding

National Natural Science Foundation of China (NSFC: 61435006, U1701661); the Research Grants Council (RGC) of Hong Kong, General Research Fund (GRF: PolyU 152658/16E, 152168/17E); The Hong Kong Polytechnic University (1-YW0S, 1-YW3G, 1-ZVFL), High-capacity fiber interconnects with optical filtering and electronic signal processing for data centers (PI, HKPU G-YBPH, 01/2017 12/2018).

Acknowledgment

The authors wish to acknowledge their gratitude to Dr. Faisal Nadeem Khan and Dr. S. Y. Zhu from the Hong Kong Polytechnic University and Dr. Feng Hao from Tianjin University for their help and fruitful discussion.

References

1. X. Huang, H. Zhang, K. Liu, T. Liu, Y. Wang, and C. Ma, "Hybrid feature extraction-based intrusion discrimination in optical fiber perimeter security system," IEEE Photonics J. **9**, 1–12 (2017).
2. J. Tejedor, J. Macias-Guarasa, H. F. Martins, D. Piote, J. Pastor-Graells, S. Martin-Lopez, P. Corredera, and M. Gonzalez-Herraez, "A novel fiber optic based surveillance system for prevention of pipeline integrity threats," Sensors **17**, 355 (2017).
3. J. K. Seedahmed S. Mahmoud, "Robust event classification for a fiber optic perimeter intrusion detection system using level crossing features and artificial neural networks," Proc. SPIE **7677**, 12 (2010).
4. S. S. Mahmoud, Y. Visagathilagar, and J. Katsifolis, "Real-time distributed fiber optic sensor for security systems: Performance, event classification and nuisance mitigation," Photonic Sensors **2**, 225–236 (2012).
5. F. Peng, H. Wu, X.-H. Jia, Y.-J. Rao, Z.-N. Wang, and Z.-P. Peng, "Ultra-long high-sensitivity ϕ-OTDR for high spatial resolution intrusion detection of pipelines," Opt. Express **22**, 13804–13810 (2014).
6. F. Peng, N. Duan, Y. Rao, and J. Li, "Real-time position and speed monitoring of trains using phase-sensitive OTDR," IEEE Photonics Technol. Lett. **26**, 2055–2057 (2014).
7. Z. Qin, L. Chen, and X. Bao, "Continuous wavelet transform for non-stationary vibration detection with phase-OTDR," Opt. express **20**, 20459–20465 (2012).
8. H. F. Martins, S. Martin-Lopez, P. Corredera, M. L. Filograno, O. F. ao, and M. González-Herráez, "Coherent noise reduction in high visibility phase-sensitive optical time domain reflectometer for distributed sensing of ultrasonic waves," J. Light. Technol. **31**, 3631–3637 (2013).

Optics EXPRESS

9. Z. Li, L. Yan, Y. Peng, W. Pan, B. Luo, and L. Shao, "Enhanced phase stability in passive analog photonic links with coherent rayleigh noise reduction," Opt. Express **23**, 5744–5748 (2015).
10. J. P. Cahill, O. Okusaga, W. Zhou, C. R. Menyuk, and G. M. Carter, "Superlinear growth of rayleigh scattering-induced intensity noise in single-mode fibers," Opt. Express **23**, 6400–6407 (2015).
11. J. Zhou, Z. Pan, Q. Ye, H. Cai, R. Qu, and Z. Fang, "Characteristics and explanations of interference fading of a ϕ-OTDR with a multi-frequency source," J. Light. Technol. **31**, 2947–2954 (2013).
12. H. Izumita, Y. Koyamada, S. Furukawa, and I. Sankawa, "Stochastic amplitude fluctuation in coherent OTDR and a new technique for its reduction by stimulating synchronous optical frequency hopping," J. Light. Technol. **15**, 267–278 (1997).
13. A. H. Hartog, *An introduction to distributed optical fibre sensors* (CRC University, Taylor & Francis Group, 2017).
14. H. Wu, S. Xiao, X. Li, Z. Wang, J. Xu, and Y. Rao, "Separation and determination of the disturbing signals in phase-sensitive optical time domain reflectometry (ϕ-OTDR)," J. Light. Technol. **33**, 3156–3162 (2015).
15. J. Tejedor, H. F. Martins, D. Piote, J. Macias-Guarasa, J. Pastor-Graells, S. Martin-Lopez, P. C. Guillén, F. D. Smet, W. Postvoll, and M. González-Herráez, "Toward prevention of pipeline integrity threats using a smart fiber-optic surveillance system," J. Light. Technol. **34**, 4445–4453 (2016).
16. K. Liu, M. Tian, T. Liu, J. Jiang, Z. Ding, Q. Chen, C. Ma, C. He, H. Hu, and X. Zhang, "A high-efficiency multiple events discrimination method in optical fiber perimeter security system," J. Light. Technol. **33**, 4885–4890 (2015).
17. Q. Sun, H. Feng, X. Yan, and Z. Zeng, "Recognition of a phase-sensitivity OTDR sensing system based on morphologic feature extraction," Sensors **15**, 15179–15197 (2015).
18. J. Tejedor, J. Macias-Guarasa, H. F. Martins, J. Pastor-Graells, S. Martín-López, P. C. Guillén, G. D. Pauw, F. D. Smet, W. Postvoll, C. H. Ahlen, and M. González-Herráez, "Real field deployment of a smart fiber-optic surveillance system for pipeline integrity threat detection: Architectural issues and blind field test results," J. Light. Technol. **36**, 1052–1062 (2018).
19. J. Tejedor, J. Macias-Guarasa, H. F. Martins, J. Pastor-Graells, P. Corredera, and S. Martin-Lopez, "Machine learning methods for pipeline surveillance systems based on distributed acoustic sensing: A review," Appl. Sci. **7**, 841 (2017).
20. M. Adeel, "Direct detected phase otdr data," https://figshare.com/s/44535a6b6888e49079dd (2018).