Distributed strain sensing by optical frequency domain reflectometry with longest common substring algorithm

Xiang Zheng,[1] Weilin Xie,[1,2,*] Qiang Yang,[1] Jiang Yang,[1] Congfan Wang,[1] Wei Wei,[1,2] and Yi Dong[1,2]
[1]Key Laboratory of Photonics Information Technology, Ministry of Industry and Information Technology, School of Optics and Photonics, Beijing Institute of Technology, No. 5, Zhongguancun South Street, Haidian District, Beijing 100081, China
[2]Yangtze Delta Region Academy of Beijing Institute of Technology, Jiaxing 314011, China
**wlxie@bit.edu.cn*

Abstract: We report a spectrum shift extraction method in optical frequency domain reflectometry based on longest common substring algorithm, allowing for an improvement in the accuracy and the range over 56% of the effective sweep range.

1. Introduction

Optical frequency-domain reflectometer (OFDR) that relies on the optical frequency-modulated continuous-wave (OFMCW) technique to interrogate the Rayleigh backscattered light (RBS) along the fiber, has been widely used in distributed fiber-optic sensing applications thanks to its high sensitivity for the demodulation of RBS changes caused by the environmental disturbances [1,2]. Spectral cross-correlation is typically applied to extract the spectrum shift between the measurement and reference spectra, thereby the physical quantities of interest. This encounters difficulties when the change of the spectrum occurs during the shift and may even fail in particular when the shift is large. Many studies have been reported to improve this extraction. Different similarity calculation methods have been compared to make a proper match for the cross-correlation [3]. A localized characteristic of the Rayleigh scattering spectrum is made use of to suppress the fake peaks and multi-peaks resulting from conventional cross-correlation [4]. The results from the cross-correlation can also be further processed with image processing methods to remove the fake peaks [5].

In this work, we propose and demonstrate a spectrum shift extraction method in OFDR relying on longest common substring (LCS) algorithm. Comparing with the conventional normalized cross-correlation, it allows for a significant enhancement in the accuracy and also permits an extension of the measurement range for the spectrum shift, thus the disturbances of interest. The successful extraction for a maximum spectrum shift accounting for 56.19% of the effective frequency sweep range has been demonstrated in distributed strain sensing.

2. Principle

In typical OFDR systems, the spectrum shift $\delta\nu$ between the reference spectrum and the measurement spectrum that are obtained via short-time Fourier transform (STFT) is usually extracted by conventional cross-correlation algorithm as shown in Fig. 1. Taking distributed strain sensing as an instance, the corresponding strain can be calculated by $\varepsilon = -\delta\nu/\kappa\nu_0$, where κ is the strain coefficient and ν_0 is the initial frequency. However, the conventional cross-correlation algorithm always has multi-peaks even fake peaks as shown in the Fig. 1(b).

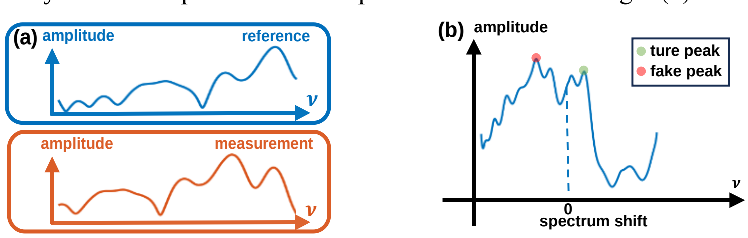

Fig. 1 (a) Typical reference and measurement spectra. (b) Cross-correlation result.

To deal with this issue, LCS algorithm that has been efficiently used to search for the longest common substring between two arbitrary strings in a two-dimension manner is introduced. The operational principle in terms of a two-dimensional matrix is illustrated in Fig. 2(a), where the raw and column of the matrix are consisted of each of the characters in the reference and targeted strings, respectively. The value for the element in raw i and column j is set to one if the corresponding j^{th} character in the targeted string is matched with the i^{th} character in the reference string, otherwise zero. This way, the longest continuous sequence of one in the diagonal (indicated by the green dashed line) is located as the longest common substring.

Extracting of the spectrum shift can be carried out in a similar manner due to the fact that analogous to a string, the spectrum can also be treated as a sequence of amplitudes at their corresponding frequencies. Inspired by this, a matching matrix with the reference and measurement spectrum acting as its raw and column is firstly constructed. Each of the element in such matrix is the relative error calculated according to Eq. (1), which represents the degree of similarity between the measured and the reference spectra at the corresponding frequencies. To simplify the extraction of the spectrum shift, a threshold is further applied to discretize the relative error and discriminate the similarity. This way, similar to Fig. 2(a), a binarized two-dimensional matrix $h(n,n)$ can be obtained with the processing indicated by Eq. (1), where n is the length, i.e. the number of spectrum samples of both the reference and the measurement spectra.

$$h(i,j)\begin{cases}0, & if\left|\frac{mea(i)-ref(j)}{ref(j)}\right| \geq \text{threshold} \\ 1, & if\left|\frac{mea(i)-ref(j)}{ref(j)}\right| < \text{threshold.}\end{cases}, i,j = 1,2,3\ldots n \tag{1}$$

where $mea(i)$ and $ref(j)$, respectively, hold for the i^{th} and the j^{th} spectrum samples in the measurement and the reference spectra. If painting the ones and zeros with white and black for each of the element, respectively, such binarized matrix can be then visualized as sketched in Fig. 2(b).

Intuitively, when no disturbance occurs, the elements in the principal diagonal should all be one, no spectrum shift is observed for the measurement spectrum with respect to the reference spectrum. In the contrary, in the presence of spectrum shift, the diagonal with all-ones will shift forwards or backwards accordingly to either side of the principal diagonal, with a shift in proportion to the amplitude of the disturbances of interest. Thus the amount of the shift for the diagonal with a continuous sequence of ones theoretically accounts for the spectrum shift.

However, it is possible that in practical, noises in the OFDR system may impair this process, resulting in a large relative error below the thresholds for the element that is supposed to be in the diagonal with a continuous sequence of ones. Therefore, we have further applied a projection towards the diagonal direction of the matrix as shown in Fig.2(b), which actually represents an accumulation for the diagonal elements of $h(n,n)$ in each of the diagonals. With a proper normalization with respect to the principal diagonal in the absence of shift, a curve $f(\delta\nu)$ describing the determination of the spectrum shift can be obtained as given in Fig. 2(b). And the spectrum shift $\delta\nu$ corresponding to the peak of $f(\delta\nu)$ is just that between the reference and measurement spectrum. Furthermore, the direction of the spectrum shift can be directly attained from the location of the peak of $f(\delta\nu)$ with respect to the position corresponding to the principal diagonal. The flow chart of the entire processing is plotted in the Fig. 2(c).

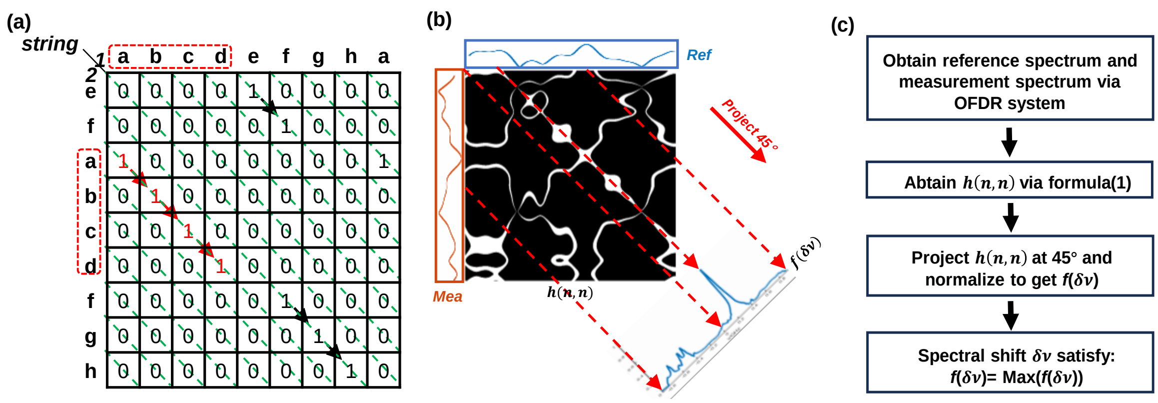

Fig. 2 (a) The principle of LCS algorithm. (b) The image of $h(n,n)$ and the spectrum shift determination curve $f(\delta\nu)$. (c) Flow diagram of OFDR distributed sensing based on LCS algorithm.

3. Experiment setup and results

The experiment setup for an OFDR system is shown in Fig. 3. The lightwave from a fiber laser is modulated through a phase modulator (PM) and subsequently filtered by a waveshaper (WS) to generate an OFMCW signal with a sweep range of 2 GHz and a sweep period of 1 ms. Then it is split into two parts through a 5:95 fiber optical coupler, acting as the reference (upper branch with 5% of the power) and the measurement (lower branch with 95% of the power) signals, respectively. The sensing part is consisted of PZT wrapped with a 60.2 m-long fiber. The beat signal is detected by a balanced photo-detector (BPD) and collected through a high-speed oscilloscope (OSC). By increasing the driving voltage applied on the PZT from 0 to 60V with an interval of 1V, the measurement spectra are obtained from the beat signal through STFT with a time window and its moving step of 200 μs and 0.5 μs, respectively. It corresponds to an effective sweep range of 1.6 GHz, leading to a spatial resolution of 25 cm. The threshold in Eq. (1) is set to 0.1.

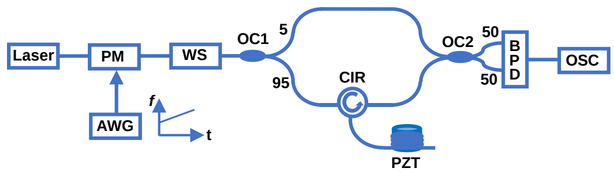

Fig. 3 Experimental setup. AWG, arbitrary waveform generator; OC, optical coupler; CIR, optical circulator.

The reference and measurement spectra at 240 kHz, corresponding to the distance of 12.28 m, is given in Fig. 4(a). With a PZT driving voltage of 14V, the resulting ~0.359 GHz spectrum shift can be spectrally identified. However, the spectrum shift derived from the normalized cross-correlation between these spectra appear to be quite misleading, in particular with a fake peak even of a higher prominence than the actual one. While a more clarified determination can be obtained by using the proposed LCS algorithm as can be inferred from the comparison between Fig. 4(b) and 4(c). It is shown that in the latter case the clearer background and higher signal-to-noise ratio leads to a more accurate estimation for the desired spectrum shift. Both algorithms are further compared in the context of distributed strain sensing by varying the driving voltage on the PZT at a 60.2 m fiber as shown by Fig. 4(d) and 4(e), respectively. The widely distributed estimation error in normalized cross-correlation could be effectively mitigated when LCS is applied. The linear fitting for the strain derived from the demodulated spectrum shift at 52.22 m by LCS algorithm exhibits an excellent fidelity with an R-square coefficient of 0.9989. It further reveals a maximum achievable strain of 5.9595 με, corresponding to a spectrum shift 0.899GHz. Considering the effective sweep range of 1.6GHz, this shift accounts for a significant enhancement for the achievable range up to 56.19% of the sweep range.

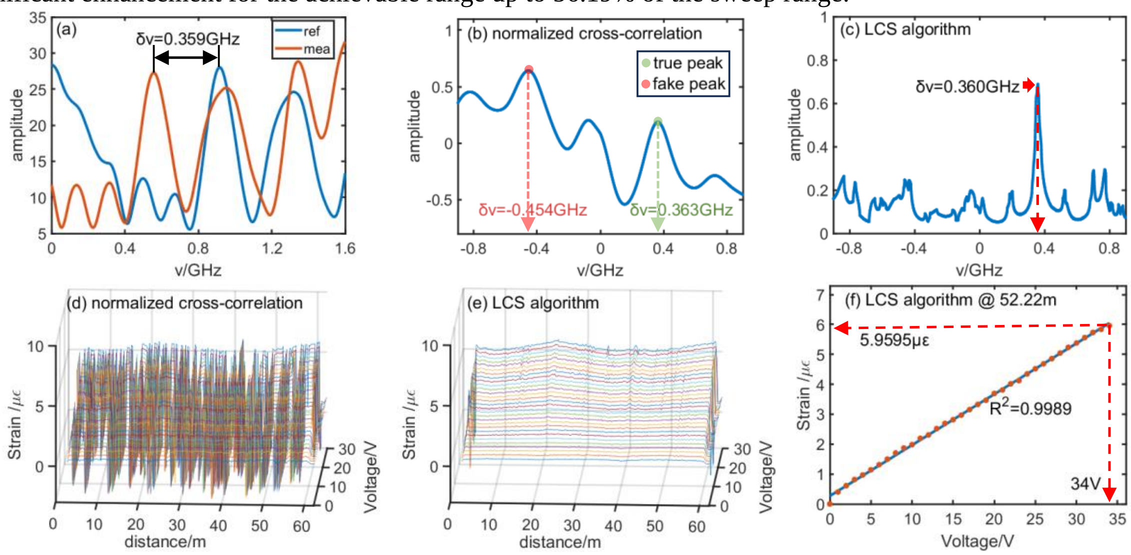

Fig. 4 (a) Reference and measurement spectrum with the PZT applied voltage of 14V @ 12.28m. (b) The normalized cross-correlation of (a). (c) The LCS algorithm result of (a). (d) Distribution strain sensing results with normalized cross-correlation. (e) Distribution strain sensing results with LCS algorithm. (f) Strain demodulation at 52.22m.

4. Conclusion

In this work, we propose a LCS based algorithm to extract the spectrum shift for the distributed sensing in OFDR. With experimental verification, compared to conventional normalized cross-correlation, the proposed method allows for significant improvement in the fidelity, accuracy, and range for the extraction of the spectrum shift.

Acknowledgement: National Natural Science Foundation of China (61827807, 62322503).

5. References

[1] Y. Koyamada, *et al*. "Fiber-optic distributed strain and temperature sensing with very high measurand resolution over long range using coherent OTDR." *J. Lightwave Technol.* **27**, 1142-1146 (2009).

[2] D. Zhou, *et al*. "Distributed vibration sensing with time-resolved optical frequency-domain reflectometry." *Opt. Express* **20**(12), 13138-13145 (2012).

[3] P. Li, *et al*. "High-spatial-resolution strain sensor based on distance compensation and image wavelet denoising method in OFDR." *J. Lightwave Technol.* **39**(19), 6334-6339 (2021).

[4] K. Feng, *et al*. "Improvement of the strain measurable range of an OFDR based on local similar characteristics of a Rayleigh scattering spectrum." *Opt. Lett.* **43**, 3293-3296 (2018).

[5] S. Qu, *et al*. "Improvement of strain measurement range via image processing methods in OFDR system." *J. Lightwave Technol.* **39**, 6340-6347 (2021).