Rayleigh-Based Distributed Optical Fiber Sensing Using Least Mean Square Similarity

Li Zhang, Zhisheng Yang, Flavien Gyger, Marcelo A. Soto† and Luc Thévenaz
EPFL Swiss Federal Institute of Technology, Institute of Electrical Engineering, SCI STI LT, Station 11, CH-1015 Lausanne, Switzerland
†Current address: Department of Electronic Engineering, Universidad Técnica Federico Santa María, 2390123 Valparaíso, Chile
Corresponding author e-mail: li.zhang@epfl.ch

Abstract: Large frequency shift errors induced by cross-correlation are investigated in Rayleigh-based distributed sensing. A least mean square method is proposed to reduce those errors, thus increasing the reliability and extending the measurand range.
OCIS codes: (060.2310) Fiber optics; (060.2370) Fiber optics sensors; (290.5870) Scattering, Rayleigh

1. Introduction

Distributed fiber sensing based on coherent Rayleigh scattering has drawn increasing attention recently due to its ultra-high sensitivity [1]. Retrieving the refractive index changes along the fiber (e.g. by any phase recovery technique), a distributed profile of the relative temperature and strain changes experienced by the sensing fiber can be obtained [2-6]. A common method to retrieve phase changes induced by refractive index variations is to cross-correlate, at each position along the fiber, the local spectral response of the coherent Rayleigh scattering acquired at a given measurement time with a reference spectrum. This spectral cross-correlation leads to a local correlation peak placed at a frequency shift that is proportional to the local refractive index change (i.e. to the local temperature and strain variations) [4-7]. However, if this frequency shift becomes larger than the width of the correlation peak, the probability of occurrence of large errors (i.e. errors larger than the correlation peak width) on the equivalent Rayleigh frequency shift (ERFS) estimation increases substantially [8], reducing the reliability of the system and limiting the measurable range of temperature/strain.

In this paper, the occurrence of large errors in the estimation of ERFS using the traditional correlation-based Rayleigh distributed sensing is investigated experimentally. An alternative method based on least mean squares (LMS) is then proposed and experimentally validated to estimate the ERFS without any cross-correlation process, thus reducing significantly the possibility of large errors and increasing the reliability of the measurements. As a proof-of-concept experiment, a 5 cm-resolution distributed fiber sensor based on phase-sensitive optical time-domain reflectometry (φ-OTDR) is implemented with direct laser modulation for frequency scanning. Experimental results show that the proposed LMS-based method increases the robustness of the φ-OTDR system, minimising the number of large errors (typically occurring when using cross-correlation), while increasing the ability of the sensor to measure large temperature changes.

2. Principle

The ERFS in both φ-OTDR and optical frequency-domain reflectometry (OFDR) is commonly estimated by cross correlation [6,7]. The spectral cross-correlation process corresponds to the inner product of two spectra (the measurement and reference spectra), whilst the resulting correlation peak is normally spectrally located where the measurement spectral response matches the one of the reference. This allows quantifying refractive index changes by estimating the ERFS of the coherent Rayleigh spectral response. However, due to this spectral shift, the measurement spectrum contains a new section that does not exist in the reference spectrum, which may give rise accidentally to a spectral positioning of the correlation peak shifted away from the 'true' ERFS. This effect has indeed been observed when using temporal cross-correlation methods to estimate relative time delays [9,10], in which they can be wrongly estimated when new data points appear into the correlated windows. The method for time delay estimation in [9,10] is mathematically fully similar to the ERFS estimation in φ-OTDR (and OFDR), and hence similar problems and limitations are to be faced in case of Rayleigh-based sensing.

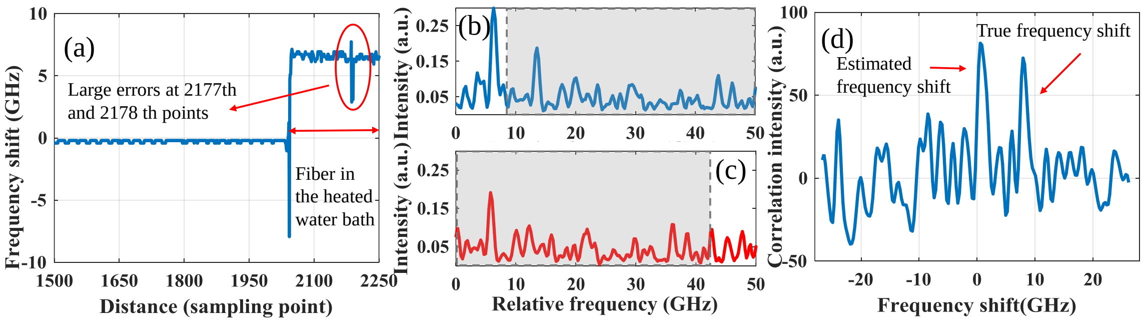

Fig. 1. An example of large frequency error with cross correlation: (a) large errors in ERFS profile (b) Reference and (c) measurement spectra at 2177th acquiring point. (d) Resulting local cross-correlation spectrum.

Figure 1(a) exemplifies a typical ERFS profile obtained using φ-OTDR associated with cross-correlation. A long hotspot corresponding to 5 GHz ERFS is implemented by deploying a section of sensing fiber in a water bath, in which large errors in the estimated ERFS appear at the 2177th and 2178th sampling points. The reference, measurement and cross-correlated spectra at the 2177th sampling point are illustrated in Fig. (b)-(d), respectively. It can be seen that, although most part of the reference spectrum in Fig. 1(b) is preserved in the measurement spectrum in Fig. 1(c) (see spectral section with grey background), the highest peak in the calculated spectral cross-correlation (see Fig. 1(d)) is not located at the 'true' ERFS but at one of the sidelobes. It can be also observed that the spectral location of this sidelobe matches the frequency difference between the highest peaks in both the reference and measurement spectra, respectively. This can be explained by the fact that the cross-correlation process corresponds to the integral of multiplied terms, which turns out to give more weight to the high-intensity points. Since the probability of the Rayleigh scattering intensity obeys an exponentially-decaying distribution [11], i.e., most spectral information is carried by low-intensity points, cross-correlated peak may be biased when high-intensity peaks appear, giving rise to large errors.

It must be pointed out that these erroneous estimations are of pure statistical nature due to the random response of coherent Rayleigh scattering and will be present even for perfect noise-free measurements.

In order to fully make use of the contribution from all spectral points with equal weight, the use of a least mean square (LMS) method is here proposed instead of cross-correlation to estimate the ERFS at each position along the sensing fiber. LMS is a common method to analyse the similarity between signals and to estimate the offset between them [9]. In this technique the difference estimator $D(\delta f)$ between the reference spectrum $S_0(f)$ and the measurement spectrum $S_1(f)$ is defined as (see Fig. 2(a) to complement the description)

$$D(\delta f) = \frac{1}{F}\int_0^F (S_1(f) - S_0(f+\delta f))^2 df, \qquad -F_0 \le \delta f \le F_0 \tag{1}$$

where δf is the frequency offset between $S_0(f)$ and $S_1(f)$ for which the similarity is tested; F is the spectral width of the measurement and F_0 is the maximum expected frequency shift, assuming that a spectral width of the reference spectrum of F+2F_0 is at least available. L and L_0 are the number of scanned frequencies within the spectral widths F and F_0, respectively, and ΔF is the frequency step, so that $F = L\Delta F$ and $F_0 = L_0\Delta F$. Then the ERFS is estimated as the δf at which the Euclidian distance $D(\delta f)$ is minimised:

$$\text{ERFS} = \delta f = \min_{\delta f \in (-F_0, F_0)} D(\delta f) \tag{2}$$

as shown in Fig. 2(b). Since the LMS-based method uses a differential operation (subtraction) to evaluate the similarity, it gives equal weight to all the points in the curve, thus reducing significantly the probability of obtaining large errors, as occuring using the conventional cross-correlation method.

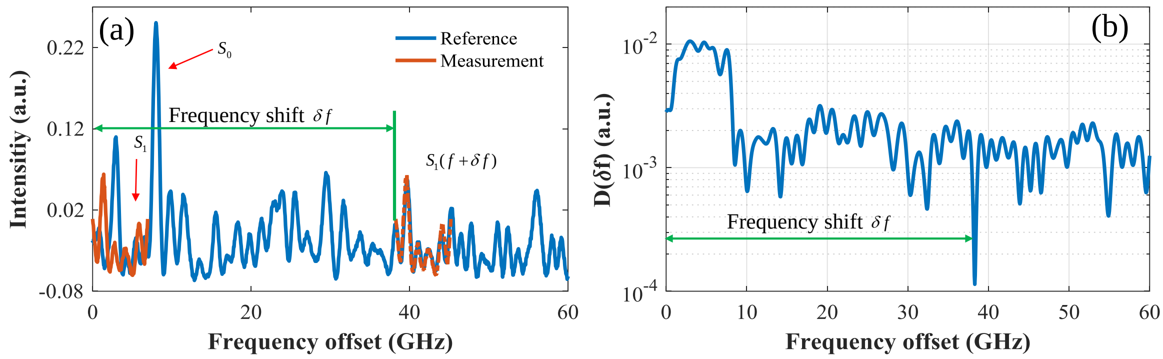

Fig. 2. Principle of the frequency shift estimation for the proposed least mean square (LMS) method. (a) The measurement spectrum at a given point (red) is swept over the broad reference spectrum (blue) and the LMS is calculated for each relative spectral position δf. (b) LMS value as a function of the frequency shift δf and the clear minimum value is the estimation of the best similarity.

3. Experimental setup

The experimental setup for validating the proposed technique is shown in Fig. 3. A distributed feedback laser with 1 MHz linewidth is used as light source. Two electro optical modulators (EOMs) are cascaded to shape the continuous wave from the laser into a pulse with high extinction ratio. Then, an erbium-doped fiber amplifier (EDFA) is appended to boost the pulse peak power (but keeping it below the onset of nonlinear effects). The pulse is then launched into the sensing fiber and its Rayleigh backscattered light is pre-amplified by another EDFA before the photo-detector (PD). The optical amplified spontaneous emission (ASE) originating from the EDFA is significantly filtered out by an optical filter with a bandwidth of 1 nm. The bandwidth of the photodetector (PD) is 3 GHz. An oscilloscope operating at a sampling rate of 5 Gs/s digitizes the electrical signal. The length of the sensing fiber is 860 m, and a 5-m hotspot near the fiber far end is immersed into a water bath in order to change the fiber temperature and to validate the measurements. The spectral response of the coherent Rayleigh scattering is obtained by sweeping the optical frequency of every interrogating optical pulse, which is achieved by directly changing the temperature/current of the laser. During the frequency scanning, each φ-OTDR trace is averaged 100 times to obtain a high enough signal-to-noise ratio (SNR), making the contribution of the noise-induced ERFS uncertainty negligible with respect to the larger statistical error on the ERFS estimation.

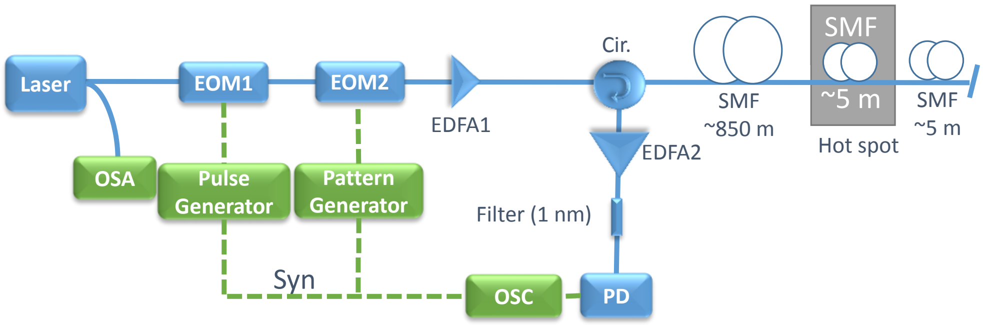

Fig. 3. Experimental setup to implement a φ-OTDR over ~860 m-long sensing fiber.(EOM: electro optical modulator; EDFA: Erbium-doped fiber amplifier, Cir: circulator; SMF: single mode fiber; PD: photodetector; OSC: oscillator; OSA: optical spectrum analyzer)

4. Results and discussions

In order to compare the probability of large errors (PLE) for both estimation methods, a set of experiments are carried out at room temperature. We firstly set the pulse width at 1 ns, corresponding to a spatial resolution of 10 cm. The frequency scanning step is set at 50 MHz here. In order to verify the impact of the number of scanning points L and the maximum expected temperature change (represented by L_0) on the PLE, ERFS estimations based on cross correlation and LMS are calculated using the same data set. It should be noted that, although the temperature has not been changed during the spectra acquisition, the number of points ($L + 2L_0$) in the reference spectrum used for the PLE calculations is larger than the respective number L in the measurement spectrum. Hence the measurement spectrum is not covering entirely the reference spectral scan, like shown in Fig.2 (a). The measurements are performed at constant temperature to minimize the impacts of other sources of errors, such as those expected from a varying temperature during the acquisition.

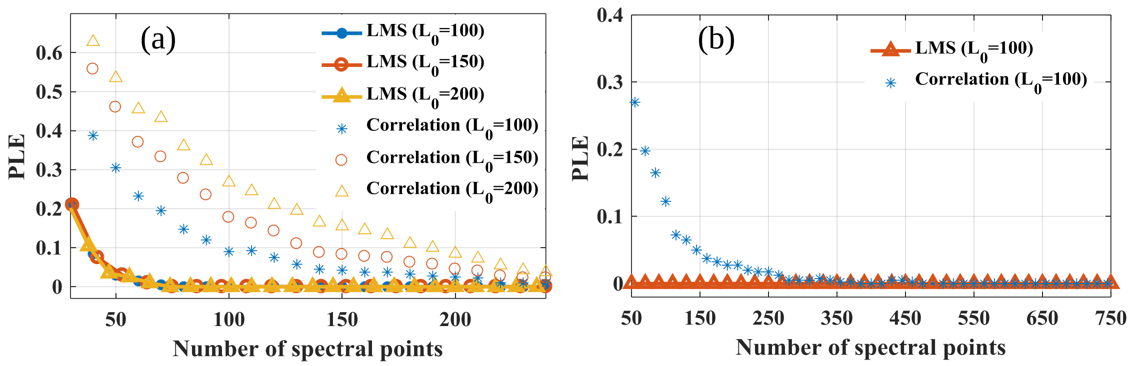

Fig. 4 Impact of the number of scanning points L on large frequency errors: (a) with noise; (b) with nullified noise

Figure 4(a) shows the PLE for both methods as a function of the number of spectral points L in the measurement spectrum. It can be observed that, compared to the cross-correlation method, the PLE is reduced significantly when using the proposed LMS and turns independent of L. To further investigate the PLE without the impact of random noise, the measurement spectra with different width are directly taken as a subset of the reference spectrum, so that the measurement spectra can perfectly match the corresponding spectral section of the reference since the noise contributions are 100% correlated. The PLEs are compared in Fig. 4(b), showing that the proposed LMS always leads to a PLE equal to zero in absence of noise and, in the case of cross-correlation, a PLE can always occur and reaches a practical value as low as 10^{-3} only when the number of measurement spectral points is larger than 450, i.e. a scanning range larger than 22.5 GHz. This demonstrates that the origin of PLE in correlation is intrinsic and unavoidable even with infinitely high SNR; however, using the proposed LMS method the PLE is solely due to the presence of noise and can nearly vanish for high SNR.

In order to validate the performance of the proposed LMS-based method, we further reduce the pulse width to 500 ps, corresponding to a spatial resolution of 5 cm. The temperature sensing experiments are carried out with a 100 MHz frequency scanning step. To enable a large measureable temperature change, a reference range extending over 120 GHz is scanned, which is achieved by tuning the laser temperature. On the other hand, according to Fig. 2(a), a large scanning range for a routine measurement is not essential once a broad reference spectrum has been acquired, so that in this case the spectral scan is realized by laser current tuning to make the acquisition faster. The scanning range of the measurement spectrum is set to cover 16 GHz. The temperature at the hotspot is varied from 0 K to 41 K over room temperature.

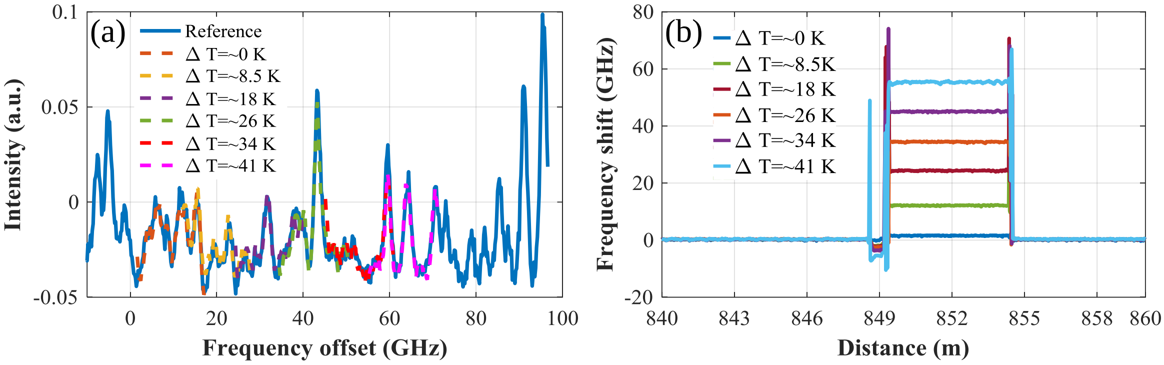

Fig. 5. Experimental results of temperature sensing using the proposed LMS-based φ-OTDR: (a) Spectra of the reference (blue line) and 6 measurements at different temperatures (at a fiber position of ~858 m); (b) Retrieved frequency shift profiles

Figure. 5(a) shows the measured spectra for different temperatures at a particular sensing point (distance ~852 m), along with the reference spectrum (blue curve). It clearly shows that the measurement spectra match well with the corresponding sections of the reference spectrum. These results also prove that the limitation of the maximum measureable temperature range is actually fixed by the spectral range of the reference and a limited spectral scanning for the running measurements brings no actual limitation on the temperature range.

The profiles of the frequency shift along the end section of the fiber are shown in Fig. 5(b), estimated with LMS at different temperatures. In the hotspot (starting from ~849.5 m), clear frequency shifts can be seen without large error, being in good agreement with the preset water temperature. A few glitches are present at fiber positions corresponding to transition to and from the heated water bath, where the temperature is obviously unstable and non-uniform. At last, the ERFS uncertainty is calculated to be 0.026 K, estimated by calculating the standard deviations of each local estimated ERFS from 10 consecutive independent measurements.

5. Conclusion

In summary, experimental results show that large errors on the frequency shift estimation of φ-OTDR can unavoidably occur when using cross-correlation techniques for pure stochastic reasons. Using the here proposed method based on least mean squares, the stochastic cause of errors can be totally suppressed and the errors are therefore significantly reduced, since only limited by noise. The maximum temperature measurement range can be safely enlarged while keeping a limited spectral scan for running acquisitions.

Acknowledgements: The work was performed in the framework of ITN-FINESSE, funded by the European Union’s Horizon 2020 research and innovation program under the Marie Sklodowska-Curie Action grant agreement n° 722509. The authors are grateful to Prof. Miguel Gonzalez Herraez and Luis Costa from the University of Alcalá for their helpful suggestions.

References

[1] H. F. Taylor and C. E. Lee, “Apparatus and method for fiber optic intrusion sensing,” U. S. Patent 5, 194-847(1993).
[2] A. Masoudi, M. Belal, and T. P. Newson, “A distributed optical fibre dynamic strain sensor based on phase-OTDR,” Meas. Sci. Technol., 24(8), 85204 (2013).
[3] Z. Wang, L. Zhang, S. Wang, N. Xue, and F. Peng , “Coherent Φ-OTDR based on I/Q demodulation and homodyne detection,” Optics Express 24(2), 853–858 (2016).
[4] J. Pastor-Graells, H. F. Martins, A. Garcia-Ruiz, S. Martin-Lopez, and M. Gonzalez-Herraez, “Single-shot distributed temperature and strain tracking using direct detection phase-sensitive OTDR with chirped pulses,” Opt. Express, 24(12), 13121-13133 (2016).
[5] S. Liehr, Y. S. Muanenda, S. Munzenberger, and K. Krebber, “Relative change measurement of physical quantities using dual-wavelength coherent OTDR,” Opt. Express, 25(2), 720-729 (2017).
[6] Y. Koyamada, M. Imahama, K. Kubota, and K. Hogari, “Fiber-optic distributed strain and temperature sensing with very high measurand resolution over long range using coherent OTDR,” J. Light. Technol., 27(9), 1142-1146 (2009).
[7] A. Sang, M. Froggatt, D. Gifford, S. Kreger, and B. Dickerson, “One centimeter spatial resolution temperature measurements in a nuclear reactor using Rayleigh scatter in optical fiber,” IEEE Sensors J., 8(7), 1375-1380 (2008).
[8] X. Bao and L. Chen, “Recent Progress in Distributed Fiber Optic Sensors,” Sensors, 12 (7), 8601-8639 (2012).
[9] G. Jacovitti and G. Scarano, “Discrete Time Techniques for Time Delay Estimation,” IEEE Trans. Signal Process, 41(2), 525-533 (1993).
[10] J. Ianniello, “Time delay estimation via cross-correlation in the presence of large estimation errors,” IEEE Trans. Acoust., 30(6), 998-1003 (1982).
[11] P. Healey, “Statistics of Rayleigh backscatter from a single-mode optical fibre,” Electronics Letters, 21(6), 226-228 (1985).