Research Article | Vol. 25, No. 25 | 11 Dec 2017 | OPTICS EXPRESS 31210
Optics EXPRESS

# Support vector machine assisted BOTDA utilizing combined Brillouin gain and phase information for enhanced sensing accuracy

HUAN WU,[1] LIANG WANG,[1,*] NAN GUO,[2] CHESTER SHU,[1] AND CHAO LU[2]

[1]*Department of Electronic Engineering, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong, China*
[2]*Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong, China*
[*]*lwang@ee.cuhk.edu.hk*

**Abstract:** Benefiting from both Brillouin amplitude and phase spectral responses during Brillouin scattering, a support vector machine (SVM) assisted Brillouin optical time domain analyzer (BOTDA) enabling the improvement of sensing accuracy with only a slight sacrifice of processing speed has been proposed and demonstrated. Only one SVM model, i.e. SVM-(g + p), is required to effectively combine the Brillouin gain and phase information in the training and testing phases, which avoids separate Brillouin gain spectrum (BGS) and Brillouin phase spectrum (BPS) fitting, and hence saves the processing time. Both simulation and experiments using different parameters were conducted to evaluate the improved performance of SVM-(g + p). Compared with the case of using BGS only or BPS only, SVM assisted BOTDA using combined BGS and BPS enhances the accuracy of temperature extraction by about 30% over a wide range of simulation and experiment parameters, only at a slight expense of the processing speed. Although the processing of both gain and phase information takes extra time, SVM-(g + p) assisted BOTDA still has a processing speed 80 times faster than that of using a conventional curve fitting method like Lorentzian curve fitting (LCF). The improved accuracy, together with fast processing speed, is crucial for future high-speed and accurate BOTDA sensors based on both Brillouin gain and phase detection.



**OCIS codes:** (060.2370) Fiber optics sensors; (290.5900) Scattering, stimulated Brillouin; (190.4370) Nonlinear optics, fibers.

## References and links

1. X. Bao and L. Chen, "Recent Progress in Distributed Fiber Optic Sensors," Sensors (Basel) **12**(7), 8601–8639 (2012).
2. L. Thévenaz, "Brillouin distributed time-domain sensing in optical fibers: state of the art and perspectives," Front. Optoelectron. **3**(1), 13–21 (2010).
3. A. Barrias, J. R. Casas, and S. Villalba, "A review of distributed optical fiber sensors for civil engineering applications," Sensors (Basel) **16**(5), 748 (2016).
4. C. Hong, Y. Zhang, G. Li, M. Zhang, and Z. Liu, "Recent progress of using Brillouin distributed fiber sensors for geotechnical health monitoring," Sens. Actuators A Phys. **258**, 131–145 (2017).
5. A. Masoudi and T. P. Newson, "Contributed Review: Distributed optical fibre dynamic strain sensing," Rev. Sci. Instrum. **87**(1), 011501 (2016).
6. Y. Dong, C. Liang, and X. Bao, "Extending the Sensing Range of Brillouin Optical Time-Domain Analysis Combining Frequency-Division Multiplexing and In-Line EDFAs," J. Lightwave Technol. **30**(8), 1161–1167 (2012).
7. F. Wang, W. Zhan, Y. Lu, Z. Yan, and X. Zhang, "Determining the Change of Brillouin Frequency Shift by Using the Similarity Matching Method," IEEE/OSA J. Lightw. Technol. **33**(19), 4101–4108 (2015).
8. H. Chang, X. Jia, X. Ji, C. Xu, L. Ao, H. Wu, Z. Wang, and W. Zhang, "DBA-Based BOTDA Using Optical Comb Pump and Pulse Coding with a Single Laser," IEEE Photonics Technol. Lett. **28**(10), 1142–1145 (2016).
9. Y. Dang, Z. Zhao, M. Tang, C. Zhao, L. Gan, S. Fu, T. Liu, W. Tong, P. P. Shum, and D. Liu, "Towards large dynamic range and ultrahigh measurement resolution in distributed fiber sensing based on multicore fiber," Opt. Express **25**(17), 20183–20193 (2017).

#306265 https://doi.org/10.1364/OE.25.031210
Journal © 2017 Received 1 Sep 2017; revised 14 Nov 2017; accepted 24 Nov 2017; published 30 Nov 2017

Optics EXPRESS

10. S. Diakaridia, Y. Pan, P. Xu, D. Zhou, B. Wang, L. Teng, Z. Lu, D. Ba, and Y. Dong, "Detecting cm-scale hot spot over 24-km-long single-mode fiber by using differential pulse pair BOTDA based on double-peak spectrum," Opt. Express **25**(15), 17727–17736 (2017).
11. A. Dominguez-Lopez, M. A. Soto, S. Martin-Lopez, L. Thevenaz, and M. Gonzalez-Herraez, "Resolving 1 million sensing points in an optimized differential time-domain Brillouin sensor," Opt. Lett. **42**(10), 1903–1906 (2017).
12. R. W. Boyd, *Nonlinear Optics*, 3rd ed. (Academic, 2007).
13. A. Zornoza, M. Sagues, and A. Loayssa, "Self-heterodyne detection for SNR improvement and distributed phase-shift measurements in BOTDA," J. Lightwave Technol. **30**(8), 1066–1072 (2012).
14. X. Tu, Q. Sun, W. Chen, M. Chen and Z. Meng, "Vector Brillouin optical time-domain analysis with heterodyne detection and IQ demodulation algorithm," IEEE Photonics J. **6**(2), 1–8 (2014).
15. X. Angulo-Vinuesa, A. Lopez-Gil, A. Dominguez-López, J. L. Cruz, M. V. Andres, S. Martin-Lopez and G. -Herraez, "Simultaneous gain and phase profile determination on an interferometric BOTDA," Proc. SPIE **9634**, 963419 (2015).
16. A. Lopez-Gil, X. Angulo-Vinuesa, A. Dominguez-Lopez, S. Martin-Lopez, and M. Gonzalez-Herraez, "Simple baseband method for the distributed analysis of Brillouin phase-shift spectra," IEEE Photonics Technol. Lett. **28**(13), 1379–1382 (2016).
17. Z. Li, L. Yan, L. Shao, W. Pan, B. Luo, J. Liang, H. He, and Y. Zhang, "Precise Brillouin gain and phase spectra measurements in coherent BOTDA sensor with phase fluctuation cancellation," Opt. Express **24**(5), 4824–4833 (2016).
18. C. Zhao, M. Tang, L. Wang, H. Wu, Z. Zhao, Y. Dang, J. Wu, S. Fu, D. Liu, and P. P. Shum, "BOTDA using channel estimation with direct-detection optical OFDM technique," Opt. Express **25**(11), 12698–12709 (2017).
19. C. Jin, L. Wang, Y. Chen, N. Guo, W. Chung, H. Au, Z. Li, H. Y. Tam, and C. Lu, "Single-measurement digital optical frequency comb based phase-detection Brillouin optical time domain analyzer," Opt. Express **25**(8), 9213–9224 (2017).
20. D. Zhou, Y. Dong, B. Wang, T. Jiang, D. Ba, P. Xu, H. Zhang, Z. Lu, and H. Li, "Slope-assisted BOTDA based on vector SBS and frequency-agile technique for wide-strain-range dynamic measurements," Opt. Express **25**(3), 1889–1902 (2017).
21. M. Niklès, L. Thévenaz, and P. A. Robert, "Brillouin gain spectrum characterization in single-mode optical fibers," J. Lightwave Technol. **15**(10), 1842–1851 (1997).
22. C. Li and Y. Li, "Fitting of Brillouin spectrum based on LabVIEW," Proc. 5th Int. Conf. Wireless Commun., Netw. Mobile Comput.,1–4 (2009).
23. C. Zhang, Y. Yang, and A. Li, "Application of Levenberg–Marquardt algorithm in the Brillouin spectrum fitting," Proc. SPIE **7129**, 71291Y (2008).
24. X. Bao, A. Brown, M. Demerchant, and J. Smith, "Characterization of the Brillouin-loss spectrum of single-mode fibers by use of very short (<10-ns) pulses," Opt. Lett. **24**(8), 510–512 (1999).
25. G. A. Ferrier, S. Afshar, X. Bao, and L. Chen, "A new fitting method for spectral characterization of Brillouin-based distributed sensors," Applications of Photonic Technology **5260**, 512–515 (2003).
26. M. A. Soto and L. Thévenaz, "Modeling and evaluating the performance of Brillouin distributed optical fiber sensors," Opt. Express **21**(25), 31347–31366 (2013).
27. A. Lopez-Gil, M. A. Soto, X. Angulo-Vinuesa, A. Dominguez-Lopez, S. Martin-Lopez, L. Thévenaz, and M. Gonzalez-Herraez, "Evaluation of the accuracy of BOTDA systems based on the phase spectral response," Opt. Express **24**(15), 17200–17214 (2016).
28. M. A. Farahani, E. Castillo-Guerra, and B. G. Colpitts, "A Detailed Evaluation of the Correlation-Based Method Used for Estimation of the Brillouin Frequency Shift in BOTDA Sensors," IEEE Sens. J. **13**(12), 4589–4598 (2013).
29. H. Wu, L. Wang, N. Guo, C. Shu, and C. Lu, "Brillouin optical time domain analyzer assisted by support vector machine for ultrafast temperature extraction," J. Lightwave Technol. **35**(19), 4159–4167 (2017).
30. C. W. Hsu and C. J. Lin, "A comparison of methods for multiclass support vector machines," IEEE Trans. Neural Netw. **13**(2), 415–425 (2002).
31. A. K. Azad, L. Wang, N. Guo, H. Y. Tam, and C. Lu, "Temperature sensing in BOTDA system by using artificial neural network," Electron. Lett. **51**(20), 1578–1580 (2015).
32. A. K. Azad, L. Wang, N. Guo, H. Y. Tam, and C. Lu, "Signal processing using artificial neural network for BOTDA sensor system," Opt. Express **24**(6), 6769–6782 (2016).

## 1. Introduction

Brillouin optical time domain analyzer (BOTDA) has attracted intensive research interest over the past three decades due to its excellent performance in distributed temperature and strain monitoring [1–5]. Since the probe is evidently amplified by the counter-propagating pump, the Brillouin frequency shift (BFS) is usually obtained by measuring the local amplitude spectral response in BOTDA, e.g. Brillouin gain spectrum (BGS) [6–11]. On the other hand, Brillouin scattering not only amplifies the probe amplitude, but also introduces a phase shift on it [12]. Thus the measurement of local phase spectral response, i.e. Brillouin

Optics EXPRESS

phase spectrum (BPS), is also an effective way of obtaining the BFS and hence the temperature or strain information [13–20]. To estimate the BFS from the measured BGS or BPS, curve fitting methods are commonly used. For BGS based BOTDA systems, as the measured BGS is the convolution between the pump pulse spectrum and the intrinsic Lorentzian gain spectrum, its lineshape can be fitted by using Lorentzian curve fitting (LCF) [21–23] or pseudo-Voigt curve fitting (pVCF) [23–25] where the BFS of the measured BGS is taken as the frequency of the peak gain on fitted curves. Quadratic curve fitting (QCF) can also be used to estimate the BFS where the measured BGS is assumed to have a quadratic lineshape near its peak [26]. While for BPS based BOTDA systems, the BFS is commonly determined by linear fitting of the BPS around the zero de-phase frequency region, as the central frequency region of the BPS in the vicinity of BFS is assumed to be quasi-linear [19, 27]. However, most of the curve fitting techniques require the initialization of model parameters and poor initialization will significantly deteriorate the accuracy in BFS determination [22, 28]. The accuracy of curve fitting techniques also depends on the number of data points collected on the measured BGS or BPS, which means fewer data points by adopting large frequency scanning step will result in quick degradation of the accuracy [27]. Moreover, due to the iterative feature, the curve fitting techniques usually require long processing time to extract BFS, especially when there are a large number of resolved sensing points along a long sensing fiber. For QCF of BGS and linear fitting of BPS, the optimal fitting spectral region in BFS estimation should be carefully determined as it affects the fitting accuracy.

To solve the drawback mentioned above in curve fitting techniques, recently we have reported a Support Vector Machine (SVM) assisted BOTDA system for ultrafast temperature extraction from the measured BGSs [29]. To extract the temperature, the measured BGSs by BOTDA are classified into different temperature classes using the SVM model after training. The corresponding temperature value of the class is taken as the measured temperature. In comparison to LCF and pVCF, SVM is more robust to a wide range of signal-to-noise ratios (SNR), averaging times, pump pulse widths, frequency scanning steps and temperatures. In addition to better accuracy of temperature extraction, SVM exhibits 100-fold faster data processing speed compared with conventional LCF and pVCF. In this paper, we propose to take full advantage of both Brillouin amplitude and phase spectral responses to further improve the accuracy of temperature extraction in SVM assisted BOTDA without much sacrifice of the processing speed. Previously, although many techniques have been proposed to measure BGS and BPS simultaneously [13–19], only BGS or BPS only is used for the sensing purpose, resulting in the waste of half of the information. Even if the two are combined for sensing in a conventional way, independent BGS fitting and BPS fitting are needed by implementing two curve fitting algorithms separately, which requires more extra time for temperature/strain extraction. Here, by effectively utilizing both Brillouin gain and phase information in a SVM assisted BOTDA, we improve the sensing accuracy by about 30% at a slight expense of the processing speed. Even with the speed reduction, the SVM assisted BOTDA using both BGS and BPS still has a processing speed 80 times faster than that of conventional LCF for temperature extraction.

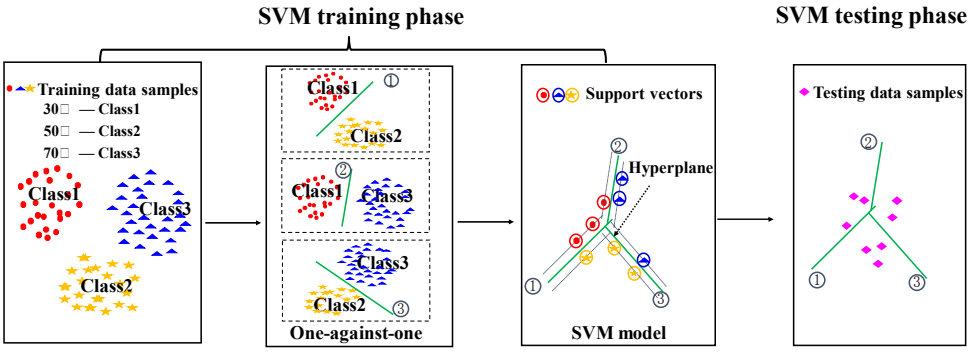
## 2. Principle and simulation

### *2.1 Principle and SVM training*

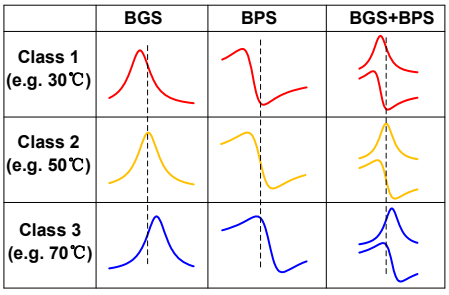
Figure. 1(a) illustrates the principle of using a linear multi-class SVM classifier for temperature extraction from the measured data samples by BOTDA. Each data sample (i.e. red dot, yellow star, blue triangle, and pink rhombus) can represent either a BGS, or a BPS, or a pair of BGS + BPS, belonging to one of the three classes. Fig. 1(b) shows the BGS only, BPS only or a pair of BGS + BPS as data samples in Class 1, Class 2, and Class 3, respectively. Here BGS and BPS are combined together to form pairs of BGS + BPS for the training of SVM and subsequent temperature extraction using SVM. Thus each data sample

Optics EXPRESS

for our proposed work is a pair of BGS + BPS, unlike the case in [29] where only BGS is used in the training and testing phase of SVM. The data points on a pair of BGS + BPS consist of the feature vector space for that data sample. The number of the data points determines the dimension of the vector space. The temperature extraction using SVM includes two phases: training and testing. In the training phase each temperature forms one class, e.g. 30°C for Class 1, 50°C for Class 2, and 70°C for Class 3, and the known pairs of BGS + BPS together with the corresponding temperature labels serve as the training data samples, as shown in Fig. 1(a). To determine the hyperplanes and support vectors of the SVM model in this multi-class classification case, 'one-against-one' strategy is adopted [29, 30], where three binary classifiers for such three temperature classes are constructed and trained to distinguish the samples of one class from another class. In the "one-against-one" process of Fig. 1(a), Hyperplane- ① is determined by the binary classifier distinguishing the samples of Class 1 from Class 2, and Hyperplane- ② and Hyperplane- ③ are determined similarly. By combining the three hyperplanes and corresponding support vectors, the SVM model for multi-class temperature classification is obtained. Finally in the testing phase the SVM model is used to classify the testing data sample (i.e. each pair of BGS + BPS measured by BOTDA) into one of the temperature classes and the corresponding temperature value of this class is regarded as the measured temperature value.




(a)




(b)

Fig. 1. (a) Principle of temperature extraction using linear multi-class SVM classifier, (b) BGSs only, BPSs only and pairs of BGS + BPS as data patterns for three different temperature classes, respectively, where the dash line indicates the BFS at a temperature of Class 2.

We use the ideal Lorentzian gain $g(\upsilon)$ and phase $p(\upsilon)$ profiles [14] as the BGS and BPS for the training of SVM:

$$g(\upsilon) = g_B \frac{1}{1+[(\upsilon-\upsilon_B)/(\Delta\upsilon_B/2)]^2} \tag{1}$$

Optics EXPRESS

$$p(\upsilon) = -\frac{2g_B \Delta\upsilon_B (\upsilon - \upsilon_B)}{\Delta\upsilon_B^2 + 4(\upsilon - \upsilon_B)^2} \tag{2}$$

where $g_B$ is the peak gain, $\upsilon_B$ is the BFS and $\Delta\upsilon_B$ is the Brillouin gain bandwidth. Each $g(\upsilon)$ and corresponding $p(\upsilon)$ form one pair of ideal BGS + BPS. 101 temperature classes are formed using the temperature range from 20°C to 70°C at a step of 0.5°C, which is enough for good accuracy. Note that the performances using temperature steps like 0.1°C and 1°C are very similar, and 0.1°C step performs only slightly better when the SNR is beyond 11dB [29]. The BFSs for ideal pairs of BGS + BPS are determined using a temperature coefficient of 1.15924 MHz/°C calibrated for our fiber under test (FUT). And for each temperature class multiple pairs of BGS + BPS with the same BFS but with bandwidth varying from 30 MHz to 100 MHz (2MHz step) are obtained to adapt Brillouin gain bandwidth variation. Finally, we have 101 × 36 training data samples to train the SVM model. The frequency range of $\upsilon$ is from 10.761GHz to 10.96GHz, the same as our frequency scanning range in the collection of BGS and BPS.

Since both Brillouin gain and phase information are used in the SVM training and testing, we use SVM-(g + p) to denote such a SVM model for convenience. In order to show the enhanced sensing accuracy by SVM-(g + p), SVM models trained only by BGS (SVM-g) or BPS (SVM-p) are also constructed for comparison. For SVM-g and SVM-p, the data samples in Fig. 1 are BGS only and BPS only, respectively. If $\upsilon$ takes 1 MHz step, there are 200 data points on each BGS and BPS as the feature vector space, and thus the dimensions of the vector space are 200 for SVM-g and 200 for SVM-p, but 400 for SVM-(g + p). In the next section, we will see that although the dimension of the vector space for SVM-(g + p) is doubled to improve the accuracy, the processing speed using SVM-(g + p) for temperature extraction only has a slight degradation. However, it is still much faster than that of the conventional curve fitting method.

### 2.2 Simulation results

In this section, we conduct simulation to analyze the performance of SVM-(g + p), and compare it with SVM-g and SVM-p. In the simulation, simulated noisy BGS and BPS are obtained by adding Gaussian white noise to the profiles from Eqs. (1) and (2), and the signal-to-noise ratio (SNR) of the simulated BGS and BPS is controlled by the amount of noise added. Then three SVM models after training, i.e. SVM-g, SVM-p, and SVM-(p + g), are employed to extract temperature from the simulated BGSs, BPSs and pairs of BGS + BPS. The simulation is run 500 times to make the results reliable for statistical analysis. The temperature to be measured is set at 60°C and the accuracy of temperature extraction is analyzed by calculating the temperature uncertainty and Root Mean Square Error (RMSE). The temperature uncertainty is defined as the standard deviation of the extracted temperature and the RMSE is calculated by comparing the temperatures given by the thermometer and extracted by SVM.

At first the Brillouin bandwidth $\Delta\upsilon_B$ is fixed at 50 MHz, corresponding to a pump pulse width around 20 ns. The step of frequency $\upsilon$ is set to be 1 MHz, corresponding to a frequency scanning step of 1 MHz. Figure 2 shows the temperature uncertainty and RMSE simulated under different SNRs by using SVM-g, SVM-p and SVM-(g + p) for temperature extraction, respectively. Here the SNR of BGS is defined as the ratio between the mean amplitude of Brillouin peak and its standard deviation, while the SNR of BPS is defined as the ratio between the mean amplitude of peak-to-peak and the standard deviation of the spectral points [26]. For all the three SVM models, the uncertainty and RMSE become worse when the SNR decreases. The uncertainty and RMSE using SVM-g and SVM-p are similar to each other, but are larger than those using SVM-(g + p), as shown by the blue triangles in Fig. 2. It implies that the accuracy of temperature extraction using SVM-(g + p) is better than that using SVM-

Optics EXPRESS

g and SVM-p. The result is due to the effective use of BGS and BPS together in the training and testing of SVM. As shown in Fig. 2, when SNR is at a relatively low level of 2.2 dB, the uncertainty and RMSE using SVM-(g + p) decrease by 31.1% and 31.2% compared with SVM-g, and by 33.8% and 33.8% compared with SVM-p. When the SNR is 14.2 dB, the uncertainty and RMSE are reduced by 30.6% and 30.3% compared with SVM-g, and by 33.1% and 32.5% compared with SVM-p. We define the improvement of accuracy as the ratio between the reduction of uncertainty/RMSE by SVM-(g + p) and uncertainty/RMSE by SVM-g or SVM-p. Compared with SVM-g and SVM-p, the average improvement of uncertainty are 29.5% and 32.7%, while that of RMSE are 29.4% and 32.6%, respectively.

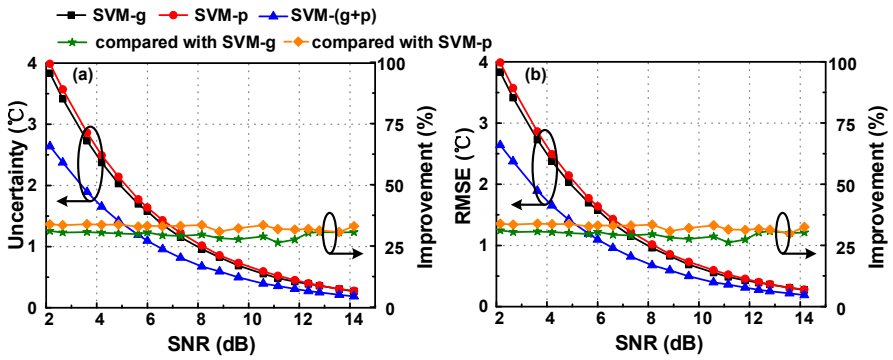



Fig. 2. (a) Temperature uncertainty, (b) root mean square error (RMSE) together with corresponding improvement simulated under different SNRs by using SVM-g, SVM-p and SVM-(g + p) for temperature extraction, respectively.

Next, we vary the Brillouin gain bandwidth from 30 MHz to 100 MHz but fix the SNR at 10.2 dB. The results are plotted in Fig. 3 from which we can see that both uncertainty and RMSE increase when the Brillouin gain bandwidth becomes large. Nevertheless, as shown by the blue triangles in Fig. 3, the uncertainty and RMSE using SVM-(g + p) are always lower than those using SVM-g and SVM-p, indicating better temperature accuracy. Compared with SVM-g, the average improvement of uncertainty and RMSE are 29.9% and 29.6%; while compared with SVM-p, the average improvement are 32.8% and 32.5%, respectively.

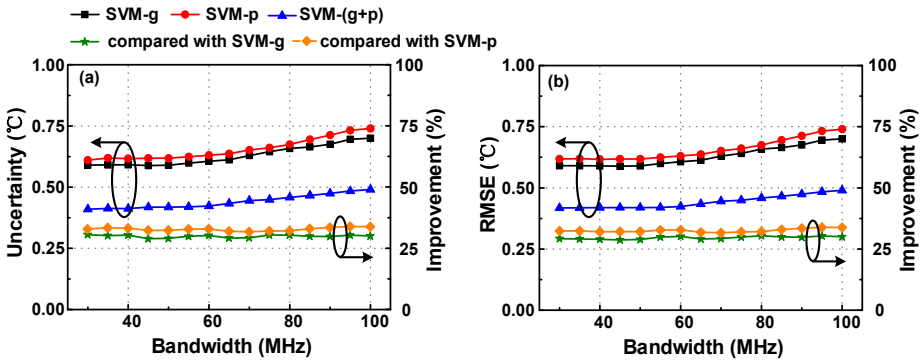



Fig. 3. (a) Temperature uncertainty, (b) root mean square error (RMSE) together with corresponding improvement simulated under different bandwidths by using SVM-g, SVM-p and SVM-(g + p) for temperature extraction, respectively.

As mentioned above, the dimension of the feature vector space for SVM-(g + p), i.e. number of data points on each pair of BGS + BPS, depends on the step of frequency $\upsilon$ . Here we also change the frequency step to examine the accuracy improvement by SVM-(g + p) under different frequency steps. The Brillouin gain bandwidth is fixed at 50 MHz and the SNR is set at 8.4 dB.

Optics EXPRESS

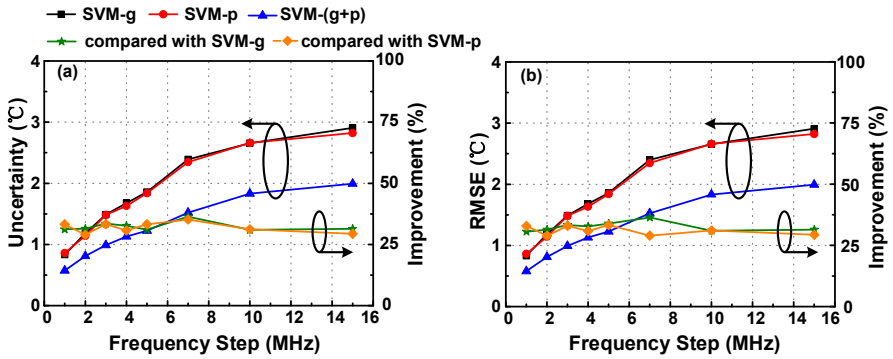



Fig. 4. (a) Temperature uncertainty, (b) root mean square error (RMSE) together with corresponding improvement simulated under different frequency scanning steps by using SVM-g, SVM-p and SVM-(g + p) for temperature extraction, respectively.

The results are given in Fig. 4. As the frequency step increases, both the uncertainty and RMSE degrade due to the reduced data points collected on each BGS, BPS and pair of BGS + BPS. At each frequency step, SVM-g and SVM-p have comparable performance, but SVM-(g + p) exhibits better accuracy, as depicted by the blue triangles in Fig. 4. In comparison to SVM-g, the improvement of uncertainty and RMSE by SVM-(g + p) are in the range of 31.0% to 36.4% and 30.6% to 36.4%; while compared with SVM-p, the corresponding improvement are in the range of 29.1% to 33.3% and 29.0% to 33.3%.

Since the gain and phase information are utilized together to improve the accuracy of temperature extraction, the dimension of the feature vector space for SVM-(g + p) is double that of SVM-g and SVM-p, which leads to a slight increase of processing time. Here we analyze the data processing time of SVM-(g + p) and compare it with that using SVM-g and SVM-p, respectively. The algorithms are implemented by using MATLAB on a conventional computer with i7-5960X CPU and 32G RAMs, and the number of training samples for SVM-g, SVM-p and SVM-(g + p) are 3636 according to Section 2.1. Table 1 shows the time needed for the training of SVM-g, SVM-p and SVM-(g + p) at different frequency scanning steps. For all the cases, the training time is less than 1s which implies the SVM training process is quite fast. To intuitively study the processing time for temperature extraction in the testing phase, SVM-(g + p) after training is employed to process 100,000 pairs of simulated BGS + BPS (Brillouin gain bandwidth ~50 MHz, SNR ~10.5 dB), equivalent to a sensing distance of 40 km fiber at a sampling rate of 250 MSample/s. Table 2 shows the processing time using SVM-(g + p) at different frequency scanning steps. For comparison, the processing times using SVM-g and SVM-p to process the same 100,000 simulated BGS and BPS are also given. We can see that compared with SVM-g and SVM-p, the processing time using SVM-(g + p) is a little larger due to the doubled dimension of the vector space by the combined BGS and BPS. However, the processing time of SVM-(g + p) is less than twice the time of SVM-g only or SVM-p only, or the sum of the time by SVM-g and SVM-p. For example, at 1MHz or 2MHz frequency step, the processing time of SVM-(g + p) is around 70% of the sum of the time by SVM-g and SVM-p; while at higher frequency steps, the processing time of SVM-(g + p) is comparable to that of SVM-g only and SVM-p only. This is because at high frequency steps, the vector space dimension difference between SVM-(g + p) and SVM-g /SVM-p becomes smaller, e.g. 20 for SVM-g, 20 for SVM-p and 40 for SVM-(g + p) at 10MHz step. On the other hand, we also calculate the processing time using conventional LCF for comparison. At 1MHz frequency step, Eq. (1) based BGS fitting and Eq. (2) based BPS fitting require 32.221 min and 41.333 min to process 100,000 simulated BGS or BPS, respectively. For the same purpose, from Table 2 we can see SVM-(g + p) only needs 22.293s to extract temperature information from 100,000 pairs of simulated BGS + BPS. Therefore, even though SVM-(g + p) slightly sacrifices the processing speed compared with SVM-g and

Optics EXPRESS

SVM-p, it is still 80-fold faster than the conventional curve fitting method and improves the accuracy of temperature extraction by around 30%.

**Table 1. Time needed for the training of SVM-g, SVM-p and SVM-(g + p)**

| Frequency step | SVM-g | SVM-p | SVM-(g + p) |
|---|---|---|---|
| **1MHz** | 0.379s | 0.380s | 0.609s |
| **2MHz** | 0.213s | 0.210s | 0.355s |
| **5MHz** | 0.120s | 0.111s | 0.143s |
| **10MHz** | 0.087s | 0.083s | 0.106s |
| **15MHz** | 0.064s | 0.057s | 0.066s |

**Table 2. Comparison of processing time using SVM-g, SVM-p and SVM-(g + p) for 100,000 sensing points**

| Frequency step | SVM-g | SVM-p | SVM-(g + p) |
|---|---|---|---|
| **1MHz** | 15.337s | 14.713s | 22.293s |
| **2MHz** | 10.295s | 10.140s | 14.693s |
| **5MHz** | 8.267s | 8.030s | 9.167s |
| **10MHz** | 8.372s | 8.029s | 8.524s |
| **15MHz** | 9.185s | 9.123s | 9.340s |

## 3. Experiment and results

In this section, we collect BGS and BPS from experiment and process them by SVM to verify the above simulation results. Here, we adopt the method in [15] to measure the BGS and BPS simultaneously along the sensing fiber. The method is relatively simple and can ensure both BGS and BPS are measured under the same conditions. In the experiment, the FUT is a 10 km long single mode fiber. The last 200 m section is free from strain and put into the oven. The sampling rate for data acquisition is 250 MSample/s. Thus, 200m fiber corresponds to 500 sampling points and provides sufficient data points for statistical analysis. Note that longer sensing range can be realized by using a 3 × 3 optical coupler structure which enables more stable measurement [16]. Coherent BOTDAs [13, 14, 17–19] can also be adopted to simultaneously measure BGSs and BPSs in a more stable manner for industrial applications. Figures 5 (a) and (b) show the BGS and BPS distribution measured with a 20ns pump pulse, 1024 times averaging and 1 MHz frequency scanning step. The measured BGSs and corresponding BPSs form pairs of BGS + BPS which are processed by the same SVM-(g + p) model in Section 2, which is trained by simulated BGSs and BPSs with different bandwidth to minimize the effect of Brillouin gain bandwidth variation along FUT on the SVM performance and also accommodate different pump pulse width [31, 32]. For comparison, the measured BGS and BPS are also processed by the same SVM-g and SVM-p as in Section 2. Figures 5(c1)-(c3) give the temperature distribution around the heated fiber section with SVM-g, SVM-p and SVM-(g + p) for temperature extraction, respectively. The temperature uncertainty/RMSE at the last 200m FUT are calculated to be 0.262°C / 0.291°C for SVM-g, and 0.285°C / 0.305°C for SVM-p; while the two parameters are improved to 0.190°C and 0.210°C by SVM-(g + p) due to the use of combined BGS and BPS. Note that the small steps in Figs. 5(c1)-(c3) are due to the fact that temperature classes are formed with a step of 0.5°C and thus the extracted temperatures by SVM have discontinuous values, e.g. 59.5°C, 60°C and 60.5°C.

Optics EXPRESS

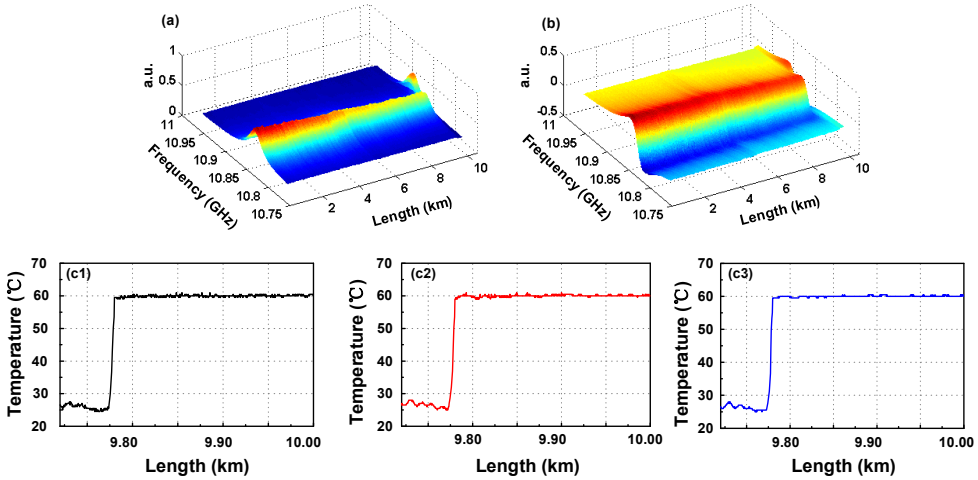



Fig. 5. (a) Measured BGS, (b) BPS distribution along 10 km FUT with last 200 m section heated at 60 °C; (c1)-(c3) temperature distribution around the heated fiber section extracted by SVM-g, SVM-p and SVM-(g + p), respectively.

Using 20 ns pump pulse and 1 MHz frequency scanning step, we then collect the BGS and BPS under different trace averaging times, i.e. 128, 256, 512, 1024. The measured SNR are calculated to be 8.4 dB, 10 dB, 11.5 dB, 13.8 dB, respectively. Figures 6(a) and (b) give the temperature uncertainty and RMSE measured under the above SNRs with SVM-g, SVM-p and SVM-(g + p) for temperature extraction, respectively. The uncertainty and RMSE are calculated using the data near the fiber end. The results are in agreement with the simulation depicted in Fig. 2 where the temperature uncertainty and RMSE by SVM-(g + p) are around 70% of those using SVM-g and SVM-p. Compared with SVM-g and SVM-p, the average improvement of uncertainty are 29.3% and 32.1%, while that of RMSE are 29.7% and 31.2%, respectively.

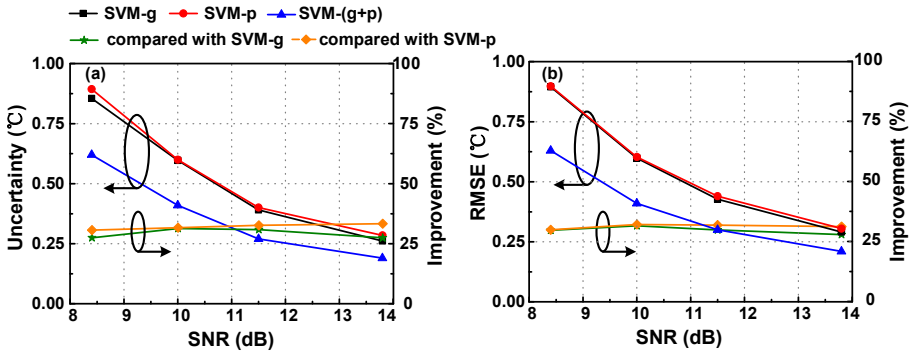



Fig. 6. (a) Temperature uncertainty, (b) RMSE together with corresponding improvement V.S. SNR measured by using SVM-g, SVM-p and SVM-(g + p) for temperature extraction, respectively.

Next, we use different pump pulse widths (i.e. 15ns, 20ns, 30ns, and 50ns) to collect the BGS and BPS under different Brillouin gain bandwidths. During the measurement, the frequency scanning step and SNR are fixed at 1 MHz and 10 dB. The corresponding gain bandwidth for the four pulse widths are measured to be 63.6MHz, 54.5MHz, 45.5MHz and 35.4MHz, respectively. Figure 7 depicts the temperature uncertainty and RMSE measured under different gain bandwidths when SVM-g, SVM-p and SVM-(g + p) are employed for temperature extraction, respectively. For all the three SVM models, a large Brillouin gain bandwidth gives rise to poor accuracy. But at each Brillouin gain bandwidth, SVM-(g + p)

Optics EXPRESS

has lower uncertainty and RMSE compared with SVM-g and SVM-p, which matches well with the simulation in Fig. 3. As an example, when the bandwidth is 63.6 MHz, the uncertainty and RMSE are 0.550°C and 0.586°C for SVM-(g + p), while they are 0.760°C and 0.783°C for SVM-g, and 0.790°C and 0.791°C for SVM-p, respectively. The average improvement of uncertainty are 29.0% and 31.4% compared with SVM-g and SVM-p; while the average improvement of RMSE are 29.2% and 30.8%, respectively.

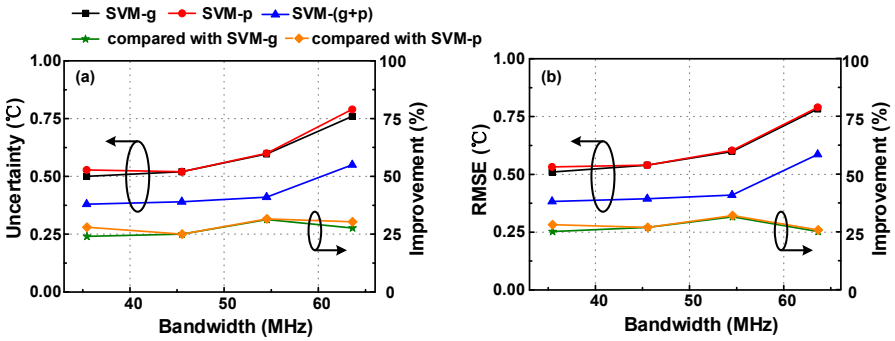



Fig. 7. (a) Temperature uncertainty, (b) RMSE together with corresponding improvement V.S. bandwidth measured by using SVM-g, SVM-p and SVM-(g + p) for temperature extraction, respectively.

Lastly, several frequency scanning steps (i.e. 1MHz, 2MHz, 5MHz, 10MHz, 15MHz) are adopted for the collection of BGS and BPS when the pump pulse width is set to be 20ns and the SNR is fixed at 8.6 dB. Figure 8 plots the measured uncertainty and RMSE versus the frequency step by using SVM-g, SVM-p and SVM-(g + p) for temperature extraction, respectively. As the frequency step increases, there are fewer data points collected on each BGS and BPS, resulting in degradation of both uncertainty and RMSE for all three SVM models. However, for each frequency scanning step, SVM-(g + p) shows better accuracy compared with the other two SVMs, which is again in agreement with the simulation in Fig. 4. The average uncertainty improvements are observed to be 31.3% and 29.1% when compared with SVM-g and SVM-p; while the average RMSE improvement ranges are 28.4% and 33.0% when compared with SVM-g and SVM-p, respectively.

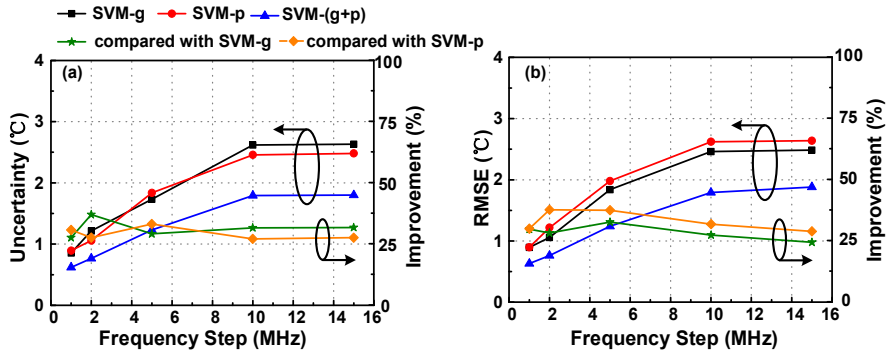



Fig. 8. (a) Temperature uncertainty, (b) RMSE together with corresponding improvement V.S. frequency step measured by using SVM-g, SVM-p and SVM-(g + p) for temperature extraction, respectively.

For the processing time of temperature extraction along the 10km FUT, Table 3 summarizes the results by using SVM-g, SVM-p and SVM-(g + p), respectively. The sampling rate for BGS and BPS acquisition is 250 MSample/s, thus 10km distance corresponds to 25,000 sensing points. Taking 1MHz frequency step as an example, SVM-g and SVM-p require 3.814s and 3.688s to process 25,000 measured BGS and BPS. In

Optics EXPRESS

comparison, SVM-(g + p) consumes 5.571s, which is only slightly longer than the previous two SVMs. If Eq. (1) based BGS fitting and Eq. (2) based BPS fitting are used for the processing of the same data, 7.916 min and 8.767min are required, respectively. Therefore, compared with SVM-g and SVM-p, SVM-(g + p) consumes only a little extra time for the accuracy improvement but it is still 80 times faster than the conventional curve fitting method.

**Table 3. Comparison of processing time for temperature extraction along 10km FUT**

| Frequency step | SVM-g | SVM-p | SVM-(g + p) |
|---|---|---|---|
| **1MHz** | 3.814s | 3.688s | 5.571s |
| **2MHz** | 2.562s | 2.533s | 3.672s |
| **5MHz** | 2.067s | 2.028s | 2.293s |
| **10MHz** | 2.091s | 2.018s | 2.133s |
| **15MHz** | 2.273s | 2.281s | 2.332s |

## 4. Conclusion

We have successfully demonstrated a SVM assisted BOTDA utilizing combined Brillouin gain and phase information for enhanced sensing accuracy. Unlike conventional curve fitting methods, where independent BGS fitting and BPS fitting are needed separately for the combination of the two information, only one SVM-(g + p) model is used. The model effectively combines the Brillouin amplitude and phase spectral responses in the training and testing phases for accuracy improvement. The performance of SVM-(g + p) is compared with SVM-g and SVM-p over a wide range of simulation and experiment parameters, achieving an accuracy improvement of about 30% due to the use of both Brillouin gain and phase information. More importantly, although the dimension of the vector space for SVM-(g + p) is doubled, the processing speed using SVM-(g + p) for temperature extraction degrades only slightly and is still 80 times faster than that using conventional LCF. The fast processing with enhanced sensing accuracy will make SVM-(g + p) a highly desirable candidate for future high-speed and high-accuracy BOTDA sensors retrieving both Brillouin gain and phase information.

## Funding

Hong Kong Research Grants Council (CUHK 14206614, 14238816, and PolyU 5208/13E); National Natural Science Foundation of China (61377093, 61435006); HK PolyU project 1-YW0S.