Research Article | Vol. 26, No. 13 | 25 Jun 2018 | OPTICS EXPRESS 16171
Optics EXPRESS

Continuous background correction using effective points selected in third-order minima segments in low-cost laser-induced breakdown spectroscopy without intensified CCD

JIANLI LIU,[1,2] RUI ZHANG,[1,2] XIAOTIAN LI,[1,5] JIANJUN CHEN,[1,2] JIANAN LIU,[1,2] JUN QIU,[1,2] XUN GAO,[3] JICHENG CUI,[1,4] AND BAYAN HESHIG[1,*]

[1]*Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun Jilin 130033, China*
[2]*University of Chinese Academy of Science, Beijing 100049, China*
[3]*Changchun University of Science and Technology, Changchun Jilin 130022, China*
[4]*jicheng_cui@163.com*
[5]*lixt_1981@163.com*
[*]*bayin888@sina.com*

Abstract: This work presents a method that can automatically estimate and remove varying continuous background emission for low-cost laser-induced breakdown spectroscopy (LIBS) without intensified CCD. The algorithm finds all third-order minima points in spectra and uses these points to partition the spectra into multiple subintervals. The mean value is then used as a threshold to select the effective points for the second-order minima in each subinterval. Finally, a linear interpolation method is used to realize extension of these effective points and complete fitting of the background using polynomials. Using simulated and real LIBS spectra with different complexities examine the validity of proposed algorithm. Additionally, five elements of five standard cast iron alloy samples are calibrated and improved very well after background removal. The results successfully prove the validity of the background correction algorithm.

OCIS codes: (300.6365) Spectroscopy, laser induced breakdown; (300.6210) Spectroscopy, atomic; (200.4560) Optical data processing.

References and links

1. C. Li, Z. Hao, Z. Zou, R. Zhou, J. Li, L. Guo, X. Li, Y. Lu, and X. Zeng, "Determinations of trace boron in superalloys and steels using laser-induced breakdown spectroscopy assisted with laser-induced fluorescence," Opt. Express **24**(8), 7850–7857 (2016).
2. D. Diaz, D. W. Hahn, and A. Molina, "Quantification of gold and silver in minerals by laser-induced breakdown spectroscopy," Spectrochim. Acta B At. Spectrosc. **136**, 106–115 (2017).
3. J. N. Kunz, D. V. Voronine, H. W. H. Lee, A. V. Sokolov, and M. O. Scully, "Rapid detection of drought stress in plants using femtosecond laser-induced breakdown spectroscopy," Opt. Express **25**(7), 7251–7262 (2017).
4. S. Chen, X. Lin, C. Yuen, S. Padmanabhan, R. W. Beuerman, and Q. Liu, "Recovery of Raman spectra with low signal-to-noise ratio using Wiener estimation," Opt. Express **22**(10), 12102–12114 (2014).
5. P. Gans and J. B. Gill, "Examination of the convolution method for numerical smoothing and differentiation of spectroscopic data in theory and in practice," Appl. Spectrosc. **37**(6), 515–520 (1983).
6. T. M. Rossi and I. M. Warner, "Fourier transform filtering of two-dimensional fluorescence data," Appl. Spectrosc. **38**(3), 422–429 (1984).
7. R. J. Berry and Y. Ozaki, "Comparision of wavelets and smoothing for denoising spectra for two-dimensional correlation spectroscopy," Appl. Spectrosc. **56**(11), 1462–1469 (2002).
8. J. Peng, S. Peng, A. Jiang, J. Wei, C. Li, and J. Tan, "Asymmetric least squares for multiple spectra baseline correction," Anal. Chim. Acta **683**(1), 63–68 (2010).
9. Z. M. Zhang, S. Chen, and Y. Z. Liang, "Baseline correction using adaptive iteratively reweighted penalized least squares," Analyst (Lond.) **135**(5), 1138–1146 (2010).
10. J. C. Cobas, M. A. Bernstein, M. Martín-Pastor, and P. G. Tahoces, "A new general-purpose fully automatic baseline-correction procedure for 1D and 2D NMR data," J. Magn. Reson. **183**(1), 145–151 (2006).

#320547 https://doi.org/10.1364/OE.26.016171
Journal © 2018 Received 24 Jan 2018; revised 4 May 2018; accepted 21 May 2018; published 12 Jun 2018

Optics EXPRESS

11. L. Shao and P. R. Griffiths, "Automatic baseline correction by wavelet transform for quantitative open-path Fourier transform infrared spectroscopy," Environ. Sci. Technol. **41**(20), 7054–7059 (2007).
12. X. H. Zou, L. B. Guo, M. Shen, X. Y. Li, Z. Q. Hao, Q. D. Zeng, Y. F. Lu, Z. M. Wang, and X. Y. Zeng, "Accuracy improvement of quantitative analysis in laser-induced breakdown spectroscopy using modified wavelet transform," Opt. Express **22**(9), 10233–10238 (2014).
13. M. S. Friedrichs, "A model-free algorithm for the removal of baseline artifacts," J. Biomol. NMR **5**(2), 147–153 (1995).
14. F. Gan, G. Ruan, and J. Mo, "Baseline correction by improved polynomial fitting with automatic threshold," Chemom. Intell. Lab. Syst. **82**(1), 59–65 (2006).
15. Z. Wang, M. Zhang, and P. B. Harrington, "Comparison of three algorithms for the baseline correction of hyphenated data objects," Anal. Chem. **86**(18), 9050–9057 (2014).
16. W. Dietrich, C. H. Rudel, and M. Neumann, "Fast and precise automatic baseline correction of one- and two-dimensional NMR spectra," J. Magn. Reson. **91**(1), 1–11 (1991).
17. S. Wartewig, *IR and Raman Spectroscopy-Fundamental Processing* (Wiley-VCH GmbH & Co. KGaA, 2013).
18. Z. Li, D. J. Zhan, J. J. Wang, J. Huang, Q. S. Xu, Z. M. Zhang, Y. B. Zheng, Y. Z. Liang, and H. Wang, "Morphological weighted penalized least squares for background correction," Analyst (Lond.) **138**(16), 4483–4492 (2013).
19. M. D. Dyar, S. Giguere, C. J. Carey, and T. Boucher, "Comparison of baseline removal methods for laser-induced breakdown spectroscopy of geological samples," Spectrochim. Acta B At. Spectrosc. **126**, 53–64 (2016).
20. I. B. Gornushkin, P. E. Eagan, A. B. Novikov, B. W. Smith, and J. D. Winefordner, "Automatic Correction of Continuum Background in Laser-induced Breakdown and Raman Spectrometry," Appl. Spectrosc. **57**(2), 197–207 (2003).
21. L. X. Sun and H. B. Yu, "Automatic estimation of varying continuum background emission in laser-induced breakdown spectroscopy," Spectrochim. Acta B At. Spectrosc. **64**(3), 278–287 (2009).
22. P. Yaroshchyk and J. E. Eberhardt, "Automatic correction of continuum background in Laser-induced Breakdown Spectroscopy using a model-free algorithm," Spectrochim. Acta B At. Spectrosc. **99**(9), 138–149 (2014).

1. Introduction

Laser-induced breakdown spectroscopy (LIBS) is a powerful analysis technology that uses a laser as a light source for excitation; complex interactions between the laser beam and materials produce specific types of radiation, and the spectral characteristics of this radiation are then used to enable material analysis [1–3]. In the initial stage of laser-induced plasma production, intense continuous background radiation overwhelms the isolated lines that represent the characteristics of the material elements. It should be noted that this continuous background is not noise but is in fact relatively stable background interference caused by bremsstrahlung radiation and recombination radiation in the plasma. The continuous background radiation decreases with increasing time and its attenuation speed is faster than that of the isolated lines, so the isolated characteristic lines that represent the elemental composition information of the sample appear gradually. In actual detection scenarios, a time-resolving detector is commonly used with the time-delay technique, in which selection of the appropriate delay time and integral time can reduce the background intensity of the spectrum signal dramatically, thus improving the signal-to-background ratio (SBR); however, it does not completely eliminate the background interference. At the same time, the laser-induced plasma is affected by laser energy instability, the matrix effect and environmental gas disturbances, while the intensity of the spectral background also changes in real time, and the spectral repeatability is thus poor. Detection and removal of the continuous background should therefore be carried out dynamically depending on the requirements of the measurement environment and the different samples. Subsequently, the intensity of the characteristic spectral lines of the LIBS spectrum used for the quantitative analysis should be the net intensity value after background removal. If the background estimation and removal process is not reasonable, then the detection performance will be limited. Therefore, accurate deduction of the continuous background has major significance for improved LIBS performance.

Background radiation and noise are problems that must be dealt with during use of LIBS spectra because both of these factors affect the accuracy of the materials analysis. Noise is a random and high-frequency signal and many methods such as the Savitzky-Golay filtering

Optics EXPRESS

method [4], convolution smoothing [5], Fourier transform methods [6], and wavelet transform methods [7] can effectively resolve the noise in the spectrum. The background radiation, however, is a more complex issue. Traditional frequency domain analysis cannot provide a theoretical description of the background. The background is well known to be a low-frequency signal with slow variation, but it is impossible to distinguish the background signal from the spectral signal on a theoretical level. It is therefore difficult to find a perfect theoretical method to solve the background problem when the real background is unknown and thus use of a background approximation is the most commonly adopted method at present.

Numerous background correction algorithms and methods have been studied and discussed in the literature, including the asymmetric least squares (ALS) method [8], the adaptive iteratively reweighted penalized least squares (Air-PLS) method [9], and the fully automatic baseline correction (FABC) method [10]; other methods include wavelet transform [11,12], median filtering [13], polynomial fitting [14,15], the iterative threshold Dietrich method [16], convex hull/rubber band techniques [17], and the morphological weighted penalized least-squares (MPLS) method [18]. All these methods have been reviewed [19] in detail by Dyar et al.

Gornushkin et al. [20] proposed an algorithm based on use of high-order polynomial functions to enable approximate fitting of the background. The algorithm finds all minima in the spectral data firstly. The spectrum is then split into a specific number of groups, denoted by N which also determines the iterative number. The minima points in each group are deemed to be "major minima". And the pixels which have intensities that fall within three standard deviations of the major minima in each group are defined as "minor minima". Next, the minimal standard average deviation σ_N between the polynomial function (drawn through major minima of all groups) and the minor minima is calculated and is used to select the optimal N. The lowest σ_N and the corresponding optimal N are obtained through multiple iterations. Finally, both the major and minor minima are used to redraw the polynomial to approximate continuum background.

The method of Gornushkin et al. [20] was improved by Sun and Yu [21]. The algorithm also begins by finding all minima in the spectral data. Then, the variance ratios of pairs of adjacent minima are calculated and are compared with a predetermined threshold to determine the minimum point to be reserved or excluded. Then a linear interpolation is performed to enlarge the set of reserved minima, and one or several polynomial functions with optimal power are finally used to draw the continuum background. Reasonable threshold selection is crucial in use of this algorithm, if the threshold is overestimated the resulting background will also be overestimated, and vice versa.

Yaroshchyk and Eberhardt [22] used a variation of a model-free algorithm proposed by Friedrichs [13] to achieve automatic correction of the continuum background in LIBS. The method is based on the idea that the baseline for a spectrum that has no analytical lines and consists only of white noise can be computed by tracking the minimum values using only the median of the noise magnitude. Use of a moving minimum for a window size W allowed the preliminary distorted baseline to be obtained. The window size is specified by the user. Finally, a normalized boxcar is convoluted with W-point moving minima to smooth off any sharp discontinuities and produce the final estimated baseline.

In this paper, we present a method for automatically correcting the varying continuous background of low-cost LIBS without intensified CCD. Our proposed method excludes the redundant points, reduces the amount of data, simplifies the spectrum and simultaneously keeps the background information intact by searching for third-order minima. The proposed method does not need to predetermine the parameters by users, thus avoiding the unfavorable effects of artificial selection parameters on background estimation. The number of segments of the spectrum is determined by the number of third-order minima points, the size of segments is determined by the interval between adjacent minimum points, and the threshold is

Optics EXPRESS

determined by the mean value of the second-order minima in the segment. What's more, the acquisition of whole spectrum effective points can be achieved by traversing all segments only once, so it needed less iteration and calculation time. The performance of our proposed method was studied by using simulated and real LIBS spectra with different complexities. Additionally, we compared the elemental calibration curves of the original spectra and the background correction spectra based on different order minima of five standard cast iron alloy samples.

2. Principle of the algorithm

2.1 Description of the algorithm

There are large numbers of local minima in the original measured spectrum that are mixed with noise and background signals, and these minimum points can be used to reflect the background information of the spectrum. However, not all of these minima are useful in the background construction process. Preservation of unnecessary minima can lead to poor background removal or excessive background removal, which leads to distortion of the spectral lines. Therefore, the key to the continuous background subtraction problem is to determine a way to filter out the effective points (where a minimum point that can be used for background fitting after screening is defined as an efficient point) across the entire spectral range. When the useful minimum points have been identified, the continuous background can then be approximated by extending it using a linear interpolation method and subsequently fitting it using polynomial functions.

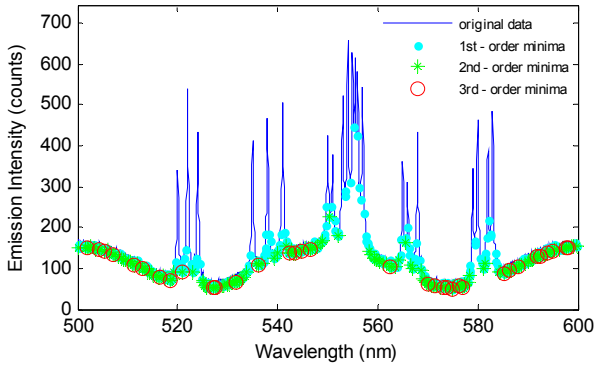

Fig. 1. Original spectrum and minimum points with different orders.

The core procedure of the algorithm is to split the spectrum into multiple subintervals by using the lower order (i.e., the third-order) minima and subsequently filter out the adjacent upper order (i.e., the second-order) minima in the interval using thresholds to obtain the effective point set required for background fitting. The spectral information can be simplified by seeking these minima points more than once. The algorithm that we propose here uses the Findpeaks function command in MATLAB software to perform the whole-spectrum multi-order minimum points search process. The command compares each data element to its neighboring values. If any data element is larger than or equal to both of its neighbors, the element is a local peak. Only the original spectral intensity is transformed into a negative number and the minimum points can thus be obtained easily.

First, the search for all the minimum intensity points is performed across the whole spectrum; these points are referred to as "first-order minima" (and are shown in Fig. 1 as cyan points) for convenience. The set of first-order minima points is then sorted according to wavelength and the search for minima is performed again; the minima points obtained in this search are called "second-order minima" (and are shown in Fig. 1 as green dots). Again, the second-order minima points set is sorted by wavelength and the minima are found again.

Optics EXPRESS

These minimum points are defined as "third-order minima" (and are shown in Fig. 1 as red dots). The minimum points for each order can be obtained sequentially by repeating the above process. The higher order minima not only contain the lower order minima, but also contain the spectral information that they carry. In the experiments, we found that the spectrum is segmented using third-order minimum points, and that the best fitting results are obtained by selecting the effective points from the second-order points in the intervals. Therefore, the algorithm uses only the second- and third-order minima points to complete estimation of the spectral background, which will be discussed in Section 2.2.

The spectrum is divided into $n + 1$ segments according to the third-order minima, which are denoted by $(\lambda_{\min 3}(g), I_{\min 3}(g))$, where $g = 1, 2, \ldots, n$. The mean value is used as the threshold to filter the second-order points in the interval with the threshold expression given by (1):

$$th(i) = m^{-1} \times \sum_{j=1}^{m} A_{ij} \tag{1}$$

where i is the interval number that is segmented using the minimum points, where $i = 1, 2, \ldots, n + 1$; m is the total number of second-order minimum points in each interval. A_{ij} represents the intensities of the second minimum points in each interval. The second minimum points in each interval are compared with the corresponding interval thresholds in sequence. The points at which the intensity exceeds the threshold are eliminated and the points with spectral intensities that are less than or equal to the threshold are retained. The set of effective points that is selected from all intervals is then used for background fitting. To prevent excessive distances between adjacent effective points from affecting the fitting results, the set of effective points is extended by linear interpolation and continuous background fitting is then realized using a high-order polynomial based on the least squares criterion. The polynomial order is determined based on the standard difference between the polynomial and the extension points, and is selected to have a value between 1 and 10.

Figure 2 shows the flow diagram for the background estimation procedure. In the first step of the procedure, the spectral data for wavelength and intensity are read to give a total of b points of the form $(\lambda(a), I(a))$. The maximum value of b is determined by the number of pixels in the charge-coupled device (CCD) detector. Subsequently, the multi-order minimum points of the original spectrum are found and the first-, second- and third-order minimum points are obtained. The original spectrum is then divided based on the third-order minimum points; there are $m = \text{locs3}(i + 1) - \text{locs3}(i) + 1$ second-order minimum points in the i-th interval, where locs3(i) represents the location of the starting minimum point in the interval. From the traversal of the ordinal number j in the i-th interval, the set of second-order minimum points of wavelength $C(i,j)$ and intensity $A(i,j)$ are obtained. The intensity threshold $E(i)$ of the i-th interval is calculated, and all the second-order minimum points $A(i,j)$ in the i-th interval are compared successively with the threshold $E(i)$. The minimum points at which the intensity exceeds the threshold is eliminated and the minimum points with intensities that are less than or equal to the threshold are written to the set $B(i,k)$. Simultaneously, the wavelength set $D(i,k)$ that corresponds with $B(i,k)$ is obtained. It is then determined whether or not the interval ordinal number i is within the range of the total number of segments $n + 1$. If i is less than $n + 1$, we then need to iterate every interval to select the effective points until i is equal or larger than $n + 1$; we then jump out of the iteration, which indicates that all effective points throughout the spectrum have been filtered.

However, the fore and aft endpoints of the adjacent intervals are repeated during the process of merging of the effective points from each interval, so these repeated effective points must be eliminated. Linear interpolation is then used to expand the selected minimum points. Finally, the high-order polynomial is used to fit the continuous background.

Optics EXPRESS

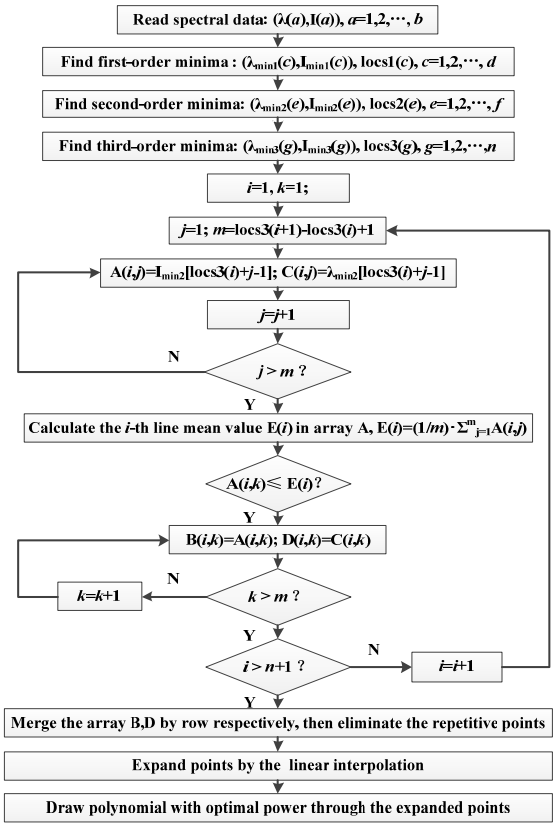

Fig. 2. Flow diagram of the continuous background estimation procedure.

2.2 Determination of the interval number

It is crucial to determine the number of intervals in the spectrum for use in the proposed algorithm. A higher number of intervals will mean that fewer minimum points are contained within each of the segmented intervals. The continuous background that is estimated based on the effective points after threshold selection will then contain too much spectral information, which then results in the occurrence of excessive deduction. A smaller number of spectral intervals will mean that more minimum points are contained within each interval. The background that is estimated based on the effective points after threshold filtering is then lower than the real spectral background, so the continuous background is not completely clean. The number of spectral intervals selected affects the continuous background removal effect directly, so it is essential to select an appropriate number of intervals to ensure the best possible removal effect. In this algorithm, the minimum points are used to segment the spectrum, and because minimum points of different orders are selected, the numbers of intervals that are generated are also different. Therefore, the problem of determination of the optimal spectral segment number is equivalent to determination of the optimal number of minimum points, and then becomes a matter of determination of which of the orders of minima is most appropriate.

Figure 3(a) shows the spectrum when segmented based on the first-order minimum points and the effective points (red dots) that are used to estimate the background are selected for the original spectrum (blue solid line) in the intervals. The original spectrum contains a large amount of information, even if the spectrum in each interval remains complex after segmented processing. This illustrates that the effective points used for screening here express too much detailed information about the spectrum, i.e., they contain a lot of spectral

Optics EXPRESS

characteristic line information, which will then lead to serious spectrum distortion after background removal.

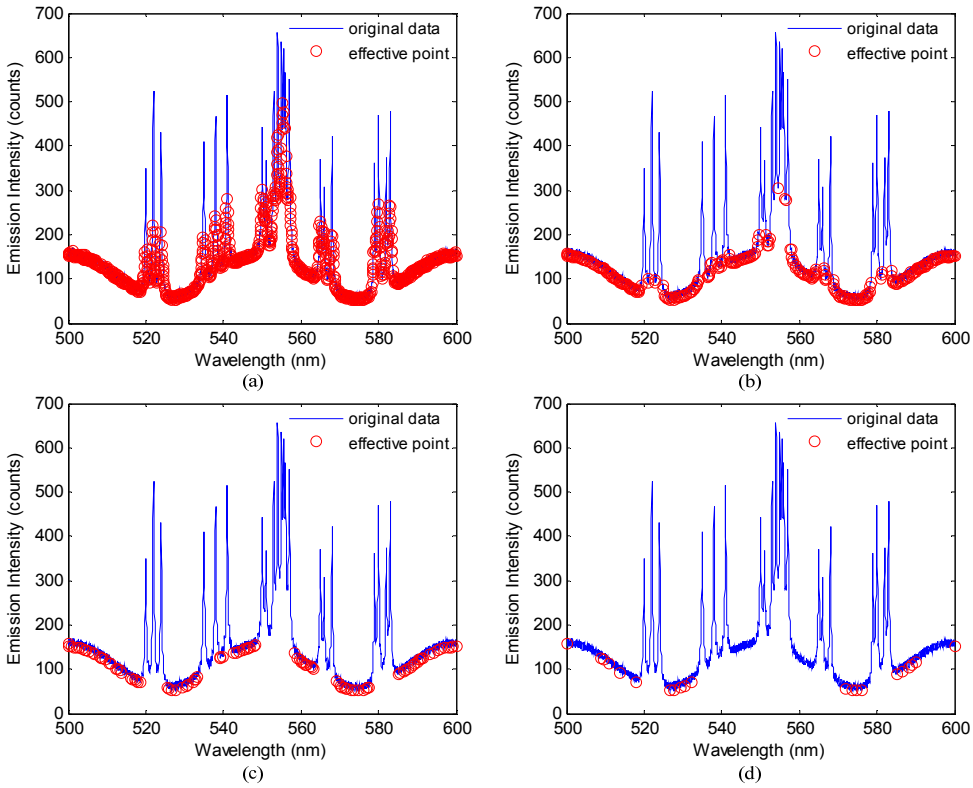

Fig. 3. Results of use of effective points selected in the (a) first-order, (b) second-order, (c) third-order, and (d) fourth-order minima segments.

Figure 3(b) shows the spectrum when segmented based on the second-order minimum points and the effective points (red dots) that are used to form the background are selected for the first-order minimum points in the segments. When compared with Fig. 3(a), the effective points in this case have obviously provided a reduction in the spectral details, but the first-order minimum points also carry more spectral characteristic information. The effective points used for filtering in the areas in which the isolated characteristic lines appear still deviate obviously from the background curve in the simulated spectrum. When the background is deducted in this case, it will also lead to overestimation and distortion of the spectral line, but the estimated background shows a significant improvement.

Figure 3(c) shows the original spectrum when segmented using the third-order minimum points and the effective points that are used to fit the background are selected for the second-order minimum points in the segmented intervals. After two attempts to find the minimum points, very little characteristic line information is contained in the second-order minimum points, but the background information is well preserved. After the redundant information in the spectrum is abandoned, the background of the selected effective points coincides fully with the cosine function of the simulated spectrum for the background. There is little deviation from the effective points at the locations of the characteristic spectral lines, which means that the isolated line information is retained completely while as much of the background information is removed as possible.

Figure 3(d) shows the spectrum when segmented on the basis of the fourth-order minimum points and the effective points that are filtered for the third-order minimum in

Optics EXPRESS

segments. After three attempts at selection, the third-order minimum points abandon almost all the isolated line information from the original spectrum, although part of the real background is also eliminated near the location of the characteristic lines at the same time. There are too few effective points for selection and the amount of information that is carried is insufficient to estimate the spectral background. When compared with Figs. 3(a) and 3(b), the background removal would be unclean, although the effective points did not deviate from the simulated background curve. This would also affect the reliability of subsequent characteristic line analysis. The other subsequent orders do not expand too much from here, but it is known from the process described above that the effect will also be reduced.

Based on the analysis above, the number of spectral intervals is determined by the third-order minimum points, and the effective points selected for the second-order minimum points in the intervals achieve the best background estimation effects. The algorithm is also suitable for use with the actual LIBS spectrum, and a correlation analysis is shown in Section 4.2.

3. Experiment setup

The experimental LIBS equipment used to obtain the spectrum used in this work is shown schematically in Fig. 4. The excitation light source is a Q-switched Nd:YAG laser (model no. Surelite III 10 from USA Continuum). The laser's output wavelength is 1064 nm, its highest pulse energy is 850 mJ, the pulse width is 8 ns, and the repetition rate is 10 Hz; it has power drift of ± 3%, a beam divergence angle of less than 0.5 mrad and a spot diameter of 9.5 mm. The experiment uses a multi-channel fiber spectrometer (model no. Avaspec-ULS2048-USB2 from Avantes). The spectral response range of the instrument is 190-557 nm and its optical resolution is 0.1 nm (full width at half maximum- FWHM). The instrument contains four 2048-pixel linear silicon CCD array detectors. Pulse triggering signal for laser and spectrometer is provided by digital delay generator (model no. DG645 from Stanford Research Systems).

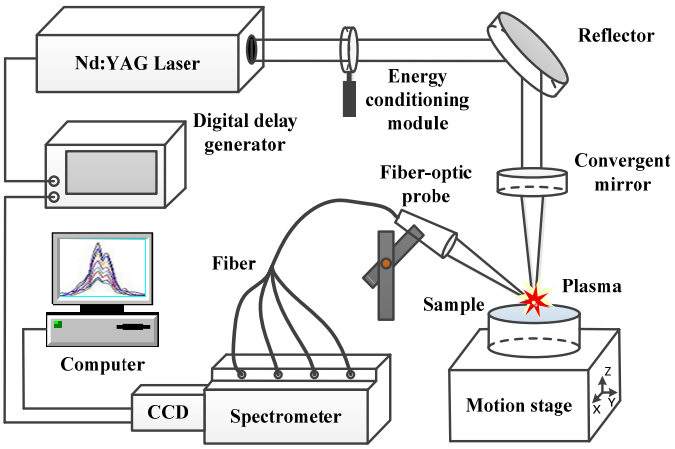

Fig. 4. Experimental LIBS system setup.

Adjusting the laser energy in the optical path by using the module composed of a gland prism and a dichroic mirror. To make full use of the experimental space, the laser beam direction is changed by 90° using a reflector, and the beam is then focused on the sample surface using a convergent mirror (focal length: 150 mm). A three-dimensional electric platform is used for sample translation during the measurement. The plasma emission is collected via a fiber-optic probe that is mounted to one side of the sample stage. The acquired LIBS spectroscopic data are then digitized and represented as two-dimensional arrays. One dimension is the ordering number of a single pixel in the range from 0 to 2047 (where each number corresponds to a specific wavelength) and the second dimension is the light intensity

Optics EXPRESS

on this pixel. The data are processed by a computer using the proposed background removal algorithm. The algorithm is based on a Matlab R2014a implementation.

4. Results and discussion

4.1 Experiments using simulated spectra

Numerous simulated spectra were generated randomly to test the proposed algorithm. Here, spectra with different complexities (i.e., simple and complex spectra) are designed to validate the algorithm. Each simulated spectrum has a range of 500-600 nm and contains a total of 2048 pixels (approximate values for single channel acquisition in a multi-channel spectrometer). The isolated characteristic spectral lines in the simulated spectra are Gaussian peaks with different intensities, and the FWHM is 0.8 nm. The curved baseline is a cosine curve and random noise is generated using a random number generator. The simulated spectra can be described using Eq. (2):

$$I(i) = I_{pure}(i) + B(i) + N(i) \tag{2}$$

where $I(i)$ is the total intensity of pixel i, $I_{\text{pure}}(i)$ is the Gaussian peak intensity, $B(i)$ is the background intensity from the cosine curve, and $N(i)$ is the white noise intensity.

4.1.1. Simple spectrum

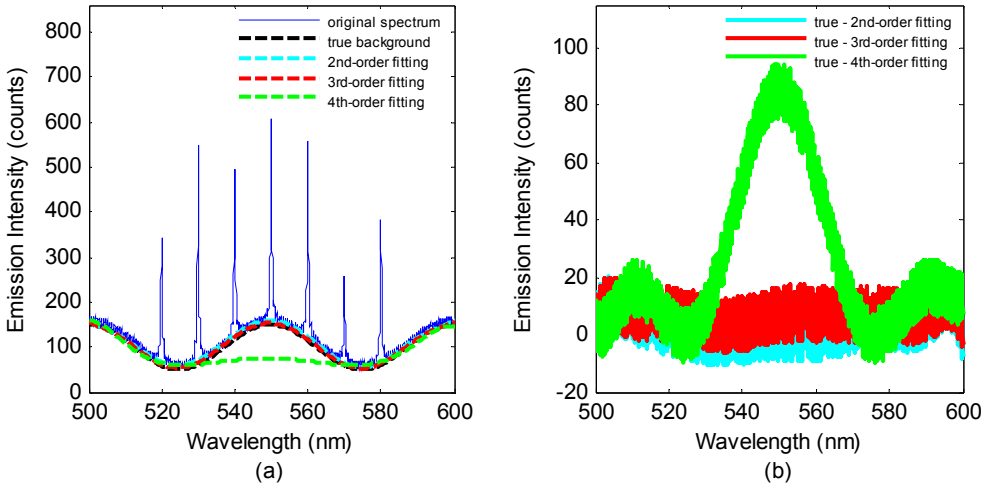

Fig. 5. Screening of effective points to fit the background (a) and the residuals (b) of the simple spectrum within the intervals of the different minimum points.

Figure 5(a) shows the simple simulated spectrum with a high signal-to-noise ratio. The simple spectrum shows seven isolated lines for analysis. Background fitting is performed using selected effective points in intervals that are segmented using minima of a different order (i.e., the second, third or fourth orders). It is obvious that the effective points that have been filtered using the fourth-order minimum points underestimate the background within the range from 530 nm to 570 nm.

Figure 5(b) illustrates the differences between the fitting background (cyan dashed line) when using the filtered effective points in the second-order minima segments and the true background (black solid line). The differences between the fitting background (red dash-dotted line) when using the effective points selected in the third-order minima segments and the true background and the differences between the background (green dotted line) estimated using the filtered effective points in the fourth-order minima segments and the true background are also shown. It can be seen from the residuals that the background that was fitted using the filtered effective points in the fourth-order minima segments differs

Optics EXPRESS

dramatically from the real background. These conditions are not sufficient for background estimation. However, the effective points that were selected in the second and third segments provided good estimates of the background. The fitting effect of the former case is such that it almost coincides with the real background, while the latter case is only slightly higher than the real background. For simple spectra, the difference between the two cases is negligible.

4.1.2. Complex spectrum

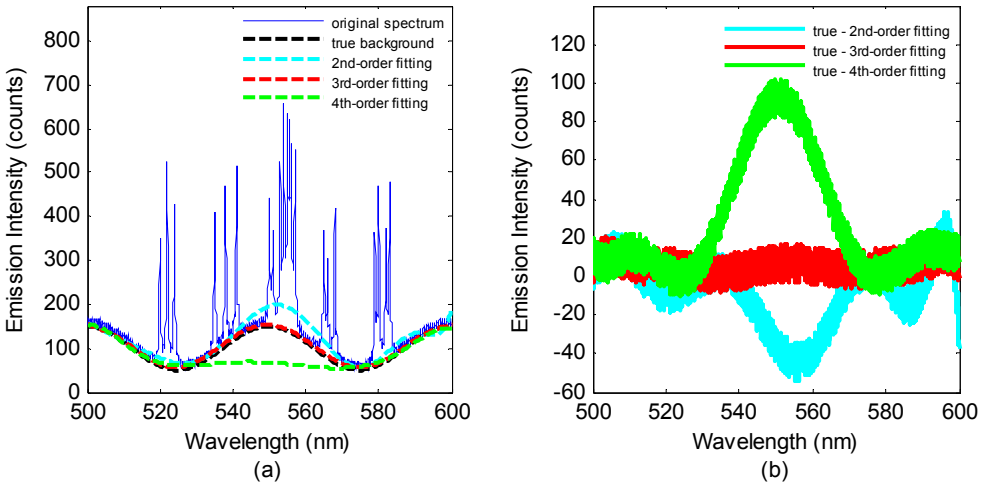

Fig. 6. Screening of effective points to fit the background (a) and the residuals (b) of the complex spectrum within the intervals of the different minimum points.

Figure 6(a) demonstrates the performance when filtered effective points are used (as shown in Figs. 3(b)-3(d)) to fit the background in minima segments of different orders. The blue line in Fig. 6(a) shows a complex spectrum with a high signal-to-noise ratio that contains 21 isolated spectral lines. The spectrum is designed to have different levels of overlap between the isolated lines, which greatly increases the difficulties of effective point selection and background fitting. It can be seen immediately from Fig. 6(a) that use of the effective points selected in the second- and fourth-order minima segments leads to failure to predict the true background of the simulated spectrum. The fitted background in the former case is overestimated across the whole spectrum, while the background estimation in the latter case is insufficient over almost the entire spectral range. In contrast, the background estimation process using effective points that were filtered in the third-order minima segments performs very well. Figure 6(b) shows the residuals between the fitted background when using effective points selected in the minima segments of the different orders and the true background. The residuals (which are the difference between the estimates when using effective points selected in the second- and fourth-order minima segments and the true background, indicated by cyan and green lines, respectively) deviate from the zero point and fluctuate greatly. The residuals of the estimation process corresponding to the use of the proposed algorithm (red line) are relatively flat with respect to the white noise intensity and are thus more useful for data processing.

4.2 Experiment using real LIBS spectra

A real LIBS spectrum has much higher complexity when compared with the simulated spectrum. To verify the validity of the proposed algorithm, we selected sparse spectra and congested spectra in different spectral ranges to perform background estimation experiments. We consider a spectrum to be either sparse or congested in terms of the number of lines and the levels of line overlaps in each individual spectrum. If a spectrum contains no isolated lines

Optics EXPRESS

(pure continuum) or only a few isolated lines, or only a few groups of lines overlap, we define that spectrum as being sparse; otherwise, we define it to be congested.

Each single LIBS spectrum collected by each channel of a multi-channel spectrometer is composed of 2048 pixels. Hundreds of LIBS spectra were generated to evaluate the algorithm, and a few examples of these spectra are shown below.

4.2.1. Sparse spectra

A series of alloy cast iron certified reference samples were used in the experiment, it is a total of six blocks (model no. NCS010001-010006-2014(T) from China NCS). Alloy cast iron samples are certified reference materials approved by the State Administration of Quality Supervision, Inspection, and Quarantine of China. The size of the NCS010003 is Ø32*28 mm, the remaining five sample sizes are Ø33*21 mm. The sample height difference will affect the results, so we use the remaining five samples for LIBS spectral detection excluding the sample of NCS010003. Figure 7 shows the sparse spectra of sample NCS010001 detected under the conditions of laser energy 5 mJ, synchronous delay 1.3 μs and integration time 1 ms.

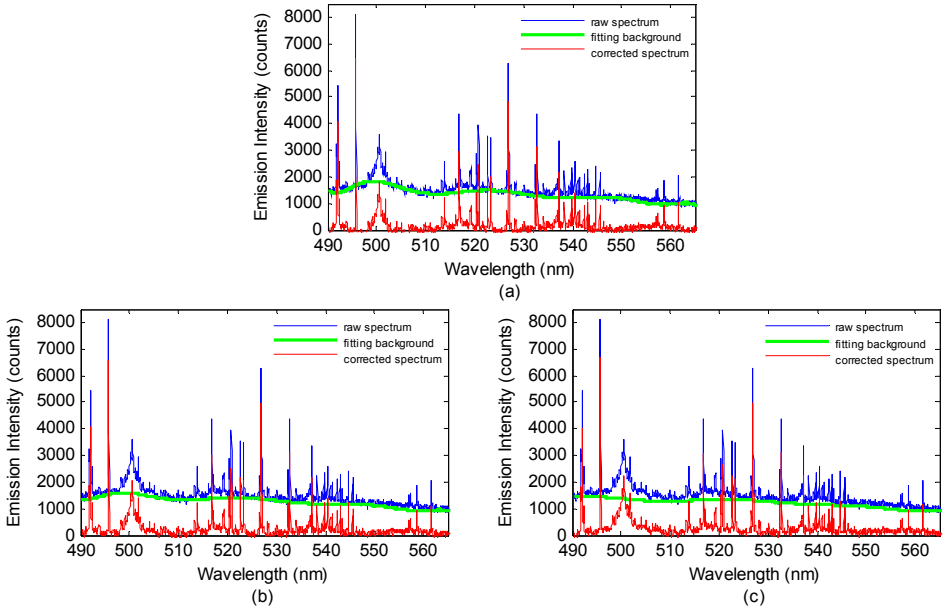

Fig. 7. Sparse spectral background fitted using effective points that were selected in the (a) second-, (b) third- and (c) fourth-order minima segments and the background-subtracted spectra.

Figure 7(a) shows that for the actual LIBS spectrum, the background that was fitted using the effective points filtered in the second-order minima segments contains some information about the characteristic spectral lines. The overestimated background appears in the region near the wavelength positions of the characteristic spectral lines and is obvious in the ranges of 497 nm-508 nm and 520 nm-530 nm. Based on the estimated background profile (green line), the spectrum after background removal (red line) has lost real information. For example, the ratio of the two feature peaks at 498 nm and 500.6 nm has changed greatly. This is equivalent to production of a new error that makes the characteristic lines lose their value for subsequent use.

Figure 7(c) deals with the same LIBS spectrum and shows that the fitted background (green line) that was formed using effective points that were selected in the fourth-order minima segments is not overestimated, but it is incomplete in the 535 nm-545 nm range.

Optics EXPRESS

Additionally, in the 494 nm-510 nm range, there is a depression in the background fitting profile that is caused by the lack of effective points in the selection process and a shortage of information in the linear interpolation and polynomial fitting stages. Background information is still present in the spectrum after background deduction (red line), which is bound to affect the spectral line analysis results.

When compared with Fig. 7(a) and 7(c), Fig. 7(b) shows that use of filtered effective points in the third-order minima segments to fit the background (green line) produces a very good performance. The estimation process tracks the spectral background profile accurately. As expected, the background estimation in the regions overlapping the spectral lines is also appropriate, and the intensity of the noncharacteristic spectral line region in the original spectrum is close to zero after background removal. The background-subtracted spectra (red line) do not show spectral distortion. The proposed algorithm thus removes as much of the background as possible while still preserving the integrity of the original characteristic line information.

4.2.2. Congested spectra

LIBS spectral detection of five-block alloy cast iron certified reference samples (model no. NCS010001, NCS010002, NCS010004, NCS010005, NCS010006) were carried out without changing the experimental conditions (laser energy 5 mJ, synchronous delay 1.3 μs and integration time 1 ms). Figure 8 shows the congested spectra of sample (model no. NCS010001) detected by the second channel of spectrometer.

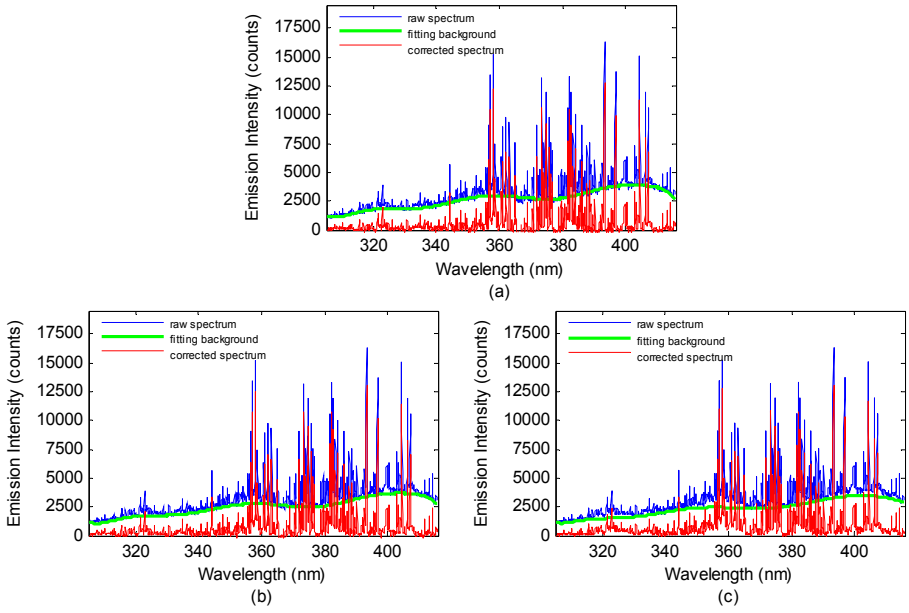

Fig. 8. Congested spectral backgrounds when fitted using effective points selected in the (a) second- (b) third- and (c) fourth-order minima segments and the background-subtracted spectra in each case.

To examine the algorithm with a more complex spectrum, a congested spectrum was collected in the 308 nm-406 nm range under non-optimal delay and integral time conditions. In this band, there are large numbers of characteristic lines of cast iron alloy elements. These lines will overlap with each other and increase the complexity of the spectrum. Therefore, it is easier to bring in minima points with deviations in screening of the effective points. Figure 8(a) shows that the background (green line) that was fitted using effective points that were filtered in the second-order minima segments describes the characteristic line in too much

Optics EXPRESS

detail. The minimum points formed between adjacent spectral lines are also selected as valid points in the spectral overlap region, which means that background overfitting appears in many regions. The background-subtracted spectrum (red line) shows that the spectral line with a wavelength of 367.1 nm is subtracted from the background and is almost drowned in the noise. The original spectral structure is also lost in the 365 nm-371 nm range. Therefore, the selection method for background prediction using the selected effective points in the second-order minima segments has failed.

From Fig. 8(c), while the fitted background (green line) has not been overestimated, it also fails to predict the true background profile in ranges between 316 nm and 330 nm, 350 nm and 367 nm, 390 nm and 413 nm, and other areas where the background removal is incomplete (too many blanks under the spectral profile). Because of the high overlap between the spectral lines in the complex spectrum, this phenomenon can be observed intuitively. As the spectral complexity increases, the defects make it more obvious that the background removal process is not clean. When compared with the original spectrum, deduction of part of the background improves the spectral line information accuracy, but the residual background intensity still affects the spectral line analysis results.

Figure 8(b) shows the spectrum when segmented using the third-order minima and the fitting background (green line) of the effective points is screened in the intervals. The proposed algorithm is thus also applicable to complex LIBS spectral fitting backgrounds and shows excellent fitting effects. The algorithm can thus remove the background to the greatest extent without losing the true characteristic spectral line information and makes up for the shortage of the other orders (such as the second or fourth orders). The algorithm has demonstrated good background estimation performances for other spectral ranges with different complexities, although these results are not described in detail here.

4.2.3. Comparison of calibrations

Table 1. Elemental compositions of five standard iron alloy samples used in calibration study

Element	Concentration (wt. %)				
	NCS010001	NCS010002	NCS010004	NCS010005	NCS010006
Mn	1.52	1.04	0.762	0.331	0.103
Cu	0.085	1.79	1.03	0.407	1.26
Ni	0.74	0.118	0.489	1.53	1.2
Cr	0.815	1.49	0.523	1.98	0.058
Mo	0.071	0.153	0.791	0.301	1.18

The real backgrounds of the experimentally collected LIBS spectra are unknown; therefore, to evaluate the accuracy of the algorithm, linear calibrations of original spectrum and background-subtracted spectrum with different order minima of five cast iron alloy samples (model no. NCS010001, NCS010002, NCS010004, NCS010005, NCS010006 from China NCS) were studied. Alloy cast iron samples are certified reference materials approved by the State Administration of Quality Supervision, Inspection, and Quarantine of China. Each sample contains the same kind of elements, each of elements in the sample is analyzed and calibrated by ICP-AES method. Table 1 lists the percentage elemental compositions of the five cast iron alloy samples. The laser energy is 5 mJ, synchronous delay is 1.3 μs and integration time is 1 ms in the experiment. Each sample was analyzed using 100 laser shots and all shots were averaged to mitigate matrix effects.

Optics EXPRESS

Table 2. Characteristics of linear calibration curves

Element				Mn	Cu	Ni	Cr	Mo
Wavelength (nm)				293.3	324.7	352.4	520.6	550.6
Calibration Curves	Untreated		*a*	2273	1536	601.8	1025	358.7
			b	3319	2370	3380	1868	1316
			R^2	0.9113	0.987	0.9722	0.9785	0.9741
	Background Corrected	Second-order Minima correct	*a*	2017	1497	600.2	1068	362.5
			b	688.3	540	616.9	337	170
			R^2	0.8959	0.9864	0.9293	0.9885	0.9597
		Third-order Minima correct	*a*	2304	1547	619.3	1068	363.6
			b	729.8	545.8	788.2	430.8	214.9
			R^2	0.9358	0.9872	0.9797	0.9903	0.9744
		Fourth-order Minima correct	*a*	2255	1487	831.1	1069	300.2
			b	890	807.9	768	487.2	298.9
			R^2	0.9602	0.9775	0.9471	0.9803	0.9689

Table 2 lists the characteristics of linear calibration curves. From the data listed in Table 2, *a* (slope), *b* (intercept) and R^2 (the goodness of fit) of the calibration curves based on the original spectra and the background-subtracted spectra were studied. The slopes of the calibration curves that were based on the original spectra and the background subtracted spectra differed and this indicates that the backgrounds of the different samples at the specific wavelengths are not the same. While the elements contained in the different samples are the same, the contents of these elements are different in each case. The effect of the matrix on the spectrum is thus different in each case, which leads to the different background estimations. This is in accordance with the changes in the slope.

The intercept of the curve is closely related to the spectral background. The continuous background of the experimentally acquired spectrum is strong and thus the original spectral calibration curve has a high intercept. When the spectral background is reduced, the intercept of the calibration curve is also reduced. But the calibration curve intercept is not equal to the spectral background, so a large intercept after the removal of the background does not represent a strong spectral background. The corrected spectrum (red line) shown in Fig. 8(b), the background does not exceed 500 at Cu 324.7 nm and 700 at Ni 352.4 nm.

For the same spectra, the effective points selected in different orders minima segments are not same, and the fitted background has differences. Therefore, the slope, intercept and R^2 of the calibration curves after background corrected have changed. As shown in Fig. 8(a), the effective points selected in the second order minima segments contain too much characteristic line information to cause the background overestimated at Cu 324.7 nm and Ni 352.4 nm, so the R^2 is decreased from 0.987 (untreated background) to 0.9864 (second-order minima correct) and from 0.9722 (untreated background) to 0.9293 (second-order minima correct), respectively. Figure 8(c) shows the effective points selected in the fourth-order minima segments are not enough to describe the real background, which causes the background removal incomplete, and the R^2 is decreased from 0.987 to 0.9775 (fourth-order minima correct) at Cu 324.7 nm and from 0.9772 to 0.9471 (fourth-order minima correct) at Ni 352.4 nm. As shown in Fig. 8(b), using the proposed algorithm to correct the background has excellent effect. This corresponds with the increase of R^2 from 0.987 to 0.9872 (third-order minima correct) at Cu 324.7 nm and from 0.9722 to 0.9797 (third-order minima correct) at Ni 352.4 nm. Similar trends for other elements (Mn, Mo) prove the reliability of the algorithm.

At Cr 520.6 nm, all the R^2 increase by using different order minima to correct the background. However, the R^2 (0.9903) is still the largest after correcting the background with proposed algorithm. The correctness of the algorithm is also illustrated.

Optics EXPRESS

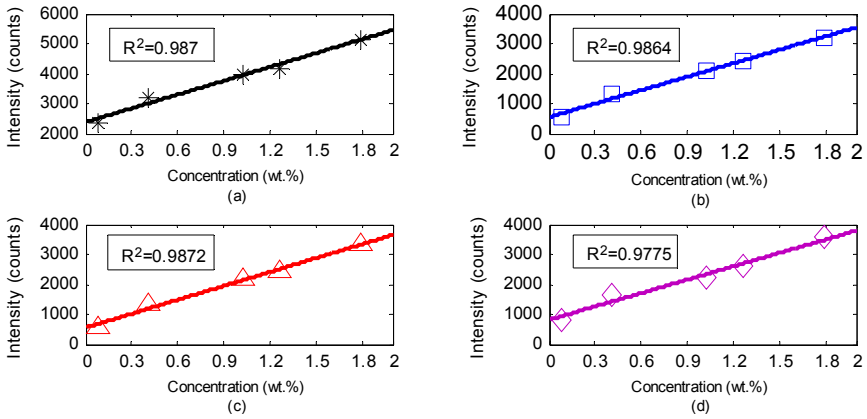

Fig. 9. Calibration curves of (a) the original data and the (b) second- (c) third- (d) fourth- order minima background subtraction data for Cu at 324.7 nm, respectively.

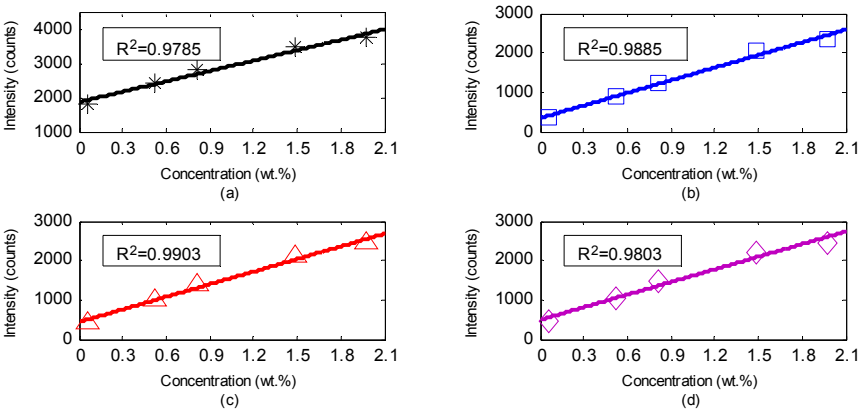

Fig. 10. Calibration curves of (a) the original data and the (b) second- (c) third- (d) fourth-order minima background subtraction data for Cr at 520.6 nm, respectively.

Figure 9 and Fig. 10 show the calibration curves of Cu (324.7 nm) and Cr (520.6 nm), respectively, based on the original spectra and the background removal spectra. The background subtraction process makes the intercepts of the calibration curves reduced a lot. The background intensity at each calibration point is different. Differences in the background deduction also make the calibration point distribution change, and the degree of fitting is thus improved. The above analysis proves that the proposed algorithm is both effective and credible.

5. Conclusions

We propose a correction algorithm that offers better performance when used in automated background correction for low-cost LIBS without intensified CCD. The algorithm does not need to set other process variables. The background is estimated by segmenting the spectra more than once and selecting effective points, and is simple to realize. Background fitting experiments are performed for both simple and complex simulated spectra using minima of different orders to cut the spectra apart and filter the effective points in the intervals. The residuals between the estimated background and the real background are then compared. The results prove that valid points (selected in the fourth-order minima segments) are not sufficient for estimation of the real background. The effective points (which were filtered in the second-order minima intervals) are successfully fitted to the simple spectral background, but produced poor predictive performance for a complex spectrum. The proposed algorithm

Optics EXPRESS

(based on screening of effective points for second-order minima in third-order minima segments) tracks the real background profile perfectly for both simple and complex spectra. To test the algorithm further, the effect comparison experiment was also performed using an actual LIBS spectrum. The algorithm can still correct varying continuum backgrounds automatically and flexibly. The background correction algorithm is suitable for use over a wide spectrum range, and an experiment to demonstrate calibration of the various elements in cast iron alloy samples also indicated the correctness of the algorithm. In conclusion, the algorithm presented in this study achieves preferable LIBS spectral background estimation and correction and provides a new concept for background removal from other types of spectra.

Funding

Chinese Finance Ministry for the National R&D Projects for Key Scientific Instruments (ZDYZ2008-1); National Natural Science Foundation of China (NSFC) (61505204); Ministry of national science and technology for National Key Basic Research Program of China (2014CB049500); National key scientific Instrument and Equipment development projects in china (2014YQ120351); Jilin Province Science & Technology Development Program Project in China (20170520167JH); Program of Jilin science and technology development (20170204051GX).