Automated image curvature assessment and correction for high-throughput Raman spectroscopy and microscopy

Ji Qi [a], Kate L. Bechtel [b] and Wei-Chuan Shih [a,*]

[a] *Department of Electrical and Computer Engineering, University of Houston, Houston, TX, USA*
[b] *Triple Ring Technologies, Newark, CA, USA*

Abstract. Raman spectroscopy and microscopy can provide molecular information for complex materials such as biological tissue and cells. In these applications, light-collection throughput is essential for speedy acquisition of high-quality data. To improve throughput, two-dimensional detectors and high numerical aperture (NA) optical systems have been employed. However, owing to the out-of-plane diffraction in grating-based dispersive spectrograph, the entrance slit image formed at the detector plane is curved along the vertical direction. Direct vertical binning of individual detector rows without correcting the curvature results in degraded spectral resolution and peak misalignment. We evaluate two software approaches to remove the image curvature after high-throughput data acquisition, with the objective to retain instrument spectral resolution and peak accuracy as if a linear-array detector were used. Curvature correction and detection are achieved in two steps: calibration of the image curvature using a Raman active material and application of the correction to future curved images. This method has been employed for a high-NA, large CCD Raman spectroscopic system deigned for non-invasive glucose sensing, a medium-NA, medium-size CCD line-scan Raman microscope designed for high-throughput tissue and cellular imaging, and an active-illumination Raman microscope. We show that remarkable improvement in data fidelity can be obtained as assessed by peak misalignment, distribution of data variance, and the waveform of principal component spectra. High quality curvature correction is essential for quantitative analysis such as the multivariate calibration, spectral pattern recognition, and peak shift detection based techniques. The software approach is highly flexible for instrument modification.

Keywords: Raman spectroscopy, Raman microscopy, SERS, curvature assessment, curvature correction

1. Introduction

Raman spectroscopy and microscopy can provide molecular information for biological tissue and cells [3,4,6,12,15,16,18–22]. In these applications, throughput is essential for speedy acquisition of high-quality data. To improve throughput, multi-channel, grating-based, dispersive spectrographs are widely employed, owing to their high efficiency and sensitivity. Unlike a monochromator, which employs an exit slit and a single-pixel detector, a multi-channel spectrograph operates without an exit slit and uses an array detector. Because Raman scattering is weak, it is important to maximize signal collection for high-quality spectra. In multi-channel spectrographs, a two-dimensional charge coupled

*Corresponding author: Wei-Chuan Shih, Department of Electrical and Computer Engineering, University of Houston, 4800 Calhoun Rd., Houston, TX 77204, USA. E-mail: wshih@uh.edu.

device (CCD) camera is often employed in conjunction with a tall grating to exploit the vertical dimension for increased throughput. Light throughput, neglecting vignetting, is proportional to the number of operational CCD pixel rows. Therefore, doubling the number of pixel rows increases the SNR by approximately 1.4 times, given a shot-noise-limited measurement. For point spectroscopy or point-scan microscopy, binning multiple CCD rows can provide higher throughput without increasing laser power or acquisition time. For microscopy applications using a line-shaped excitation pattern, highly parallel acquisition from equivalently 100 spots along the line pattern can effectively boost imaging throughput by ∼100X compared to a point-scan system with the same power density. However, given the combination of high numerical aperture (NA) and a large area CCD, a problem arises – slit image curvature – due to out-of-plane diffraction [23]. If vertical binning is applied directly, the resulting spectral resolution is degraded with peak misalignment, introducing errors in subsequent peak assignment and quantitative multivariate analysis. In the line-scan microscopy case, the peak misalignment prevents direct and precise assembly of the final three-dimensional data cube (i.e., x, y and λ) for subsequent multivariate analysis.

Various hardware approaches, such as curved slits or convex spherical gratings, have been pursued [7,23]. In the curved slit approaches, fiber bundles have been employed as shape transformers to increase Raman light collection efficiency. At the entrance port the fibers are arranged in a round shape to accommodate the focal spot, and at the exit port arranged in a curved line, to counteract the curvature distortion. This exit arrangement serves as the entrance slit of the spectrograph, and provides immediate first order correction of the curved image, as described below. Such first-order correction may not be satisfactory with the presence of substantial change of the image curvature across the wavelength range of interest (∼150 nm) and narrow spectral features.

As an alternative to the hardware approach, software can be employed to remove fluorescence backgrounds [9] and filter out cosmic rays [8,17], and correct the curved image, with potentially better accuracy and flexibility for system modification. For example, Pelletier et al. corrected infrared absorption spectra using a software technique [11]. Esmonde-White et al. reported a projective transformation technique to correct curvature and image distortions [5]. Based on similar principles, we have developed our own method involving the use of a highly Raman active reference material to provide a sharp image on the CCD. Using the curvature of the slit image at the center wavelength as a guide, we determine by how many pixels in the horizontal direction each off-center CCD row needs to be shifted in order to generate a linear vertical image. This method, similar to the curved-fiber-bundle hardware approach, does not take the wavelength-dependent slit curvature into account. The resulting spectral quality of the "shifting" method is thus equivalent to the curved-fiber-bundle hardware approach [23], as evidenced by comparison of our respective results. This issue is exacerbated when large CCD chips and high-NA spectrographs are employed for increasing the throughput of the Raman scattered light. For example, for non-invasive glucose sensing, our system combines an $f/1.4$ spectrograph and a 1 inch^2 CCD chip. We show that a modified software technique can significantly reduce the image curvature.

Another recent trend of utilizing large CCD chips can be found in state-of-the-art parallel Raman microscopes employing either a line-shaped laser or active-illumination with semi-arbitrary spot patterns, where curvature assessment and correction is essential. Therefore, we have tested our method on a home-built line-scan system for various samples such as tissue, polymer microparticles, drugs and self-assembled monolayers. Our home-built system has an $f/2$ spectrograph coupled to an 8×26 mm^2 CCD chip. We show that remarkable improvement in data fidelity can be obtained as assessed by peak misalignment, distribution of data variance, and the waveform of principal component spectra.

2. Methods

2.1. Theory and simulations

The spectrograph is modeled as a $4f$ optical system with an entrance slit at the front focal point of the first lens, a grating at the conjugate Fourier plane, and a CCD at the back focal point of the second lens. The image shape of the first order diffracted light can be modeled using the following equation [23]:

$$\frac{x_{\text{CCD}}}{f} = \frac{\lambda \cdot y_{\text{CCD}}^2}{2p \cdot \cos\beta_0 \cdot f}. \tag{1}$$

In Eq. (1), x_{CCD} and y_{CCD} are the CCD coordinates, f is the focal length of the lens before the CCD, p is the pitch of the diffraction grating, λ is the wavelength, and β_0 the diffraction angle of a given wavelength, for example, the center wavelength (905 nm) of the spectral range of our system (830 nm $<$ $\lambda <$ 970 nm). We note that the curvature is a function of wavelength, the only variable in addition to the CCD coordinates, x_{CCD} and y_{CCD}. To simulate the response of our instrument, we use Eq. (1) and actual parameters from our instrument: The CCD dimensions 2.68 cm (H) $\times$ 2.6 cm (V) with 20×20 μm^2 pixel size; the focal length 8.5 cm; and the spectral range 830–970 nm with 830 nm laser excitation.

The impulse response of the system for an infinitesimally narrow entrance slit is plotted in Fig. 1(a) for five representative wavelengths. To better examine the wavelength dependence of the image curvature, each of the plotted curves is shifted to the 905 nm line with their vertexes aligned (Fig. 1(b)). Without correction, vertical binning results in a highly degraded spectrum with resolution $\sim$35 pixels ($\sim$54 cm^{-1}) (full width half maximum (FWHM)), or equivalently 0.05 pixel per vertical pixel shift. The first-order correction method (referred to subsequently as *Method* 1) of simply shifting CCD rows assuming the curvature remains constant over the entire spectral range results in the curves of Fig. 1(c). The uncorrected error is as large as $\pm$15 pixels ($\sim \pm$23 cm^{-1}) at both ends of the CCD using Method 1, causing significant linewidth broadening.

2.2. Experimental section

In the first example, we employ the fiber bundle approach with the exit fibers arranged in a straight line (as opposed to the curved exit end approach described earlier). The fiber bundle (RoMack Fiber
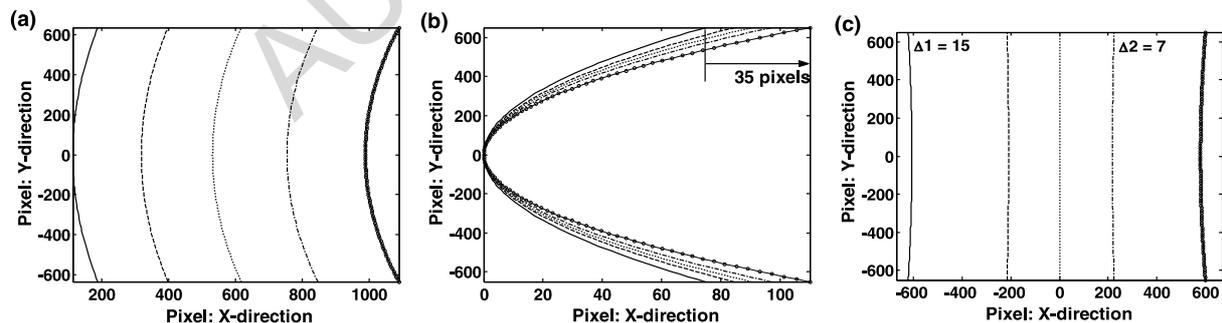
Fig. 1. (a) Simulated impulse response of the system at 5 different wavelengths for an infinitesimally narrow slit. The CCD is 1340 (H) $\times$ 1300 (V) pixels with 20 μm^2 pixel size. "–": 830 nm, "--": 880 nm, "$\cdots$": 905 nm, "----": 930 nm, "Φ": 970 nm. (b) Curves in (a) shifted such that their apexes are aligned and with the x-axis expanded to show detail. The largest difference is 35 pixels if the whole CCD range is used. (c) After the first-order curvature correction with Method 1. The uncorrected error is still approximately 15 pixels on either side of the CCD.

Optics) consists of 65 cladding-stripped fibers with 400 µm core diameter. The linear shape at the exit end serves as the entrance slit, with equivalent dimensions 0.4 mm (H) × 26 mm (V), and is imaged ∼1.1× onto the a liquid nitrogen cooled CCD (1340 × 1300 pixels, Spec-10, Princeton Instruments). A Kaiser HoloSpec $f/1.4$ spectrometer (Kaiser Optical Systems) was modified to incorporate the fiber bundle with the collimating stage removed.

In the line-scan Raman microscope, a 133 µm × 1 µm line-shaped laser illumination is projected onto the sample with the Raman scattered light from the entire line imaged with 60× to the entrance slit (100 µm (H) × 8 mm (V)) of an $f/2$ spectrograph (LS-785, Princeton Instruments) with a thermo-electrically cooled CCD (400 × 1300 pixels, PIXIS 400BR, Princeton Instruments) [14].

Method 1 calibrates the image curvature at one wavelength and uses this information to shift off-center rows correspondingly. The fact that the curvature increases towards the higher dispersion end is not taken into account. Furthermore, we found that due to the curvature change across the spectral range of interest, each row spectrum appears to be "stretched" differently compared to the center-row spectrum. As a result, the same spectral coverage occupies different number of pixels in each row, i.e., the center-row spectrum has the fewest number of pixels whereas the top- or bottom-row the most.

Our improved method (subsequently referred to as Method 2) employs the following scheme. We first calibrate several spectral lines (where peaks in a reference sample occur) and use these as boundaries to separate the row spectra into several spectral segments. The chosen peaks are then aligned to their respective locations in the center-row spectrum. Linear interpolation is then used to "compress" each row spectrum back to the same length as the center-row spectrum for each segment, while maintaining signal conservation. The resulting row spectra are well aligned and can be used directly as a line image or vertically binned to obtain a single spectrum.

We have implemented the method using a two-step approach: Image curvature calibration and correction. For calibration, a full-frame image is taken with a reference material that has prominent peaks across the spectral range of interest, for example, acetaminophen powder. We chose several prominent peaks across the wavelength range of interest, as depicted by the arrows in Fig. 2. The calibration algorithm then generates a map of the amount of shift for each CCD pixel and a scale factor to maintain signal conservation in each CCD row. Once the map and the scale factor are generated, usually when the system is modified or requires re-calibration, the correction algorithm can be applied to future measurements. A key advantage of the two-step approach is that curve fitting is only needed in the calibration step, not the application step. Excellent curve fitting quality can be obtained from high SNR images taken in the calibration step. We integrated the algorithm written in MATLAB (The MathWorks) with LabVIEW (National Instruments) data acquisition software to streamline data processing.

3. Results and discussion

We have implemented both software curvature correction methods. The raw and corrected experimental full-frame Raman spectra of acetaminophen powder (Sigma-Aldrich) before software vertical binning are shown in Fig. 3. Method 1 corrects the majority of the curvature, however, there is substantial improvement from Method 1 to Method 2, especially toward either side of the CCD, as can be seen by comparing Fig. 3(c) and (e). The overall linewidth reduction in 14 prominent peaks is 7% (FWHM). This is a substantial improvement given that the spectral resolution is ∼15 cm^{-1} owing to the large (∼400 µm) slit width. If a narrower slit is employed for better spectral resolution, the overall linewidth reduction will be more pronounced. Note that the images were taken with 5-pixel CCD hardware vertical
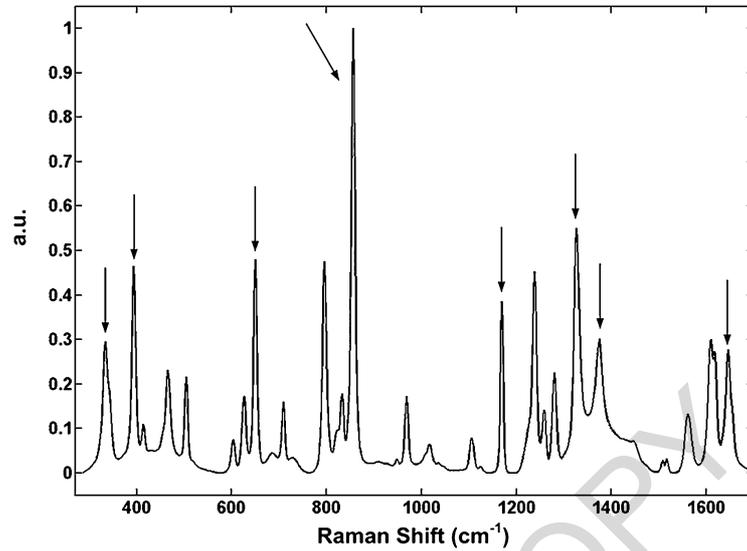
Fig. 2. Raman spectrum of acetaminophen powder, used as the reference material in the calibration step. Eight prominent peaks used as separation boundaries are indicated by arrows.
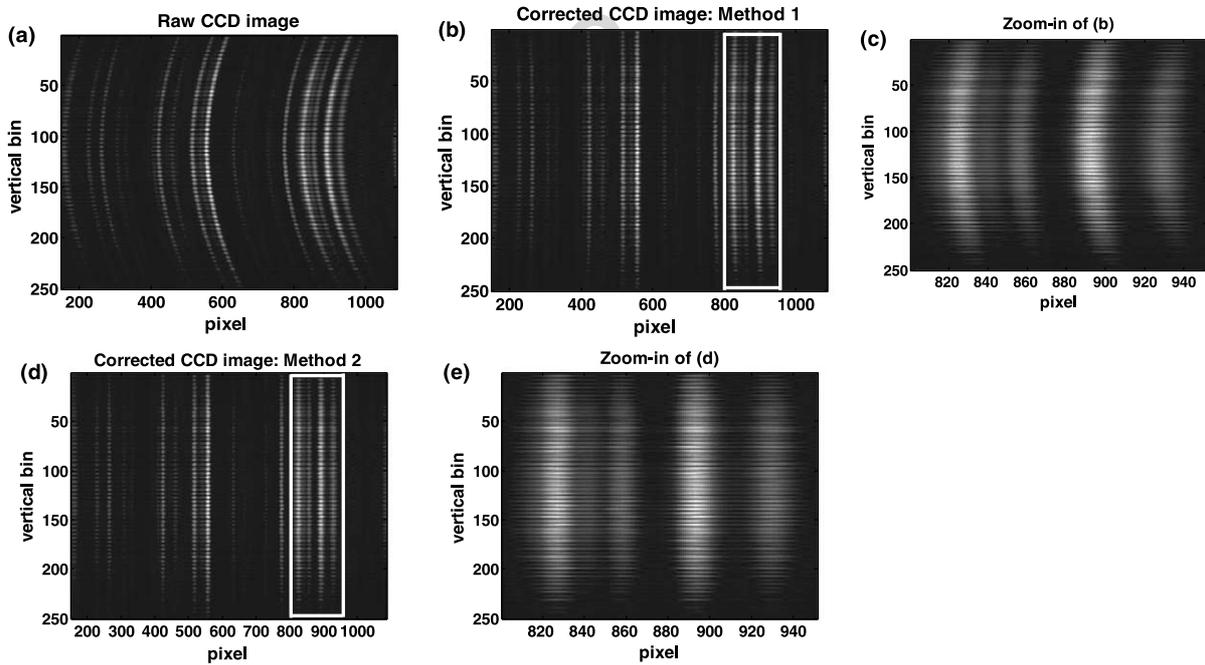
Fig. 3. CCD image of acetaminophen powder. Images were created with 5-pixel hardware binning. (a) Raw image; (b) after applying Method 1; (c) zoom-in of the box in (b); (d) after applying Method 2; (e) zoom-in of the box in (d).
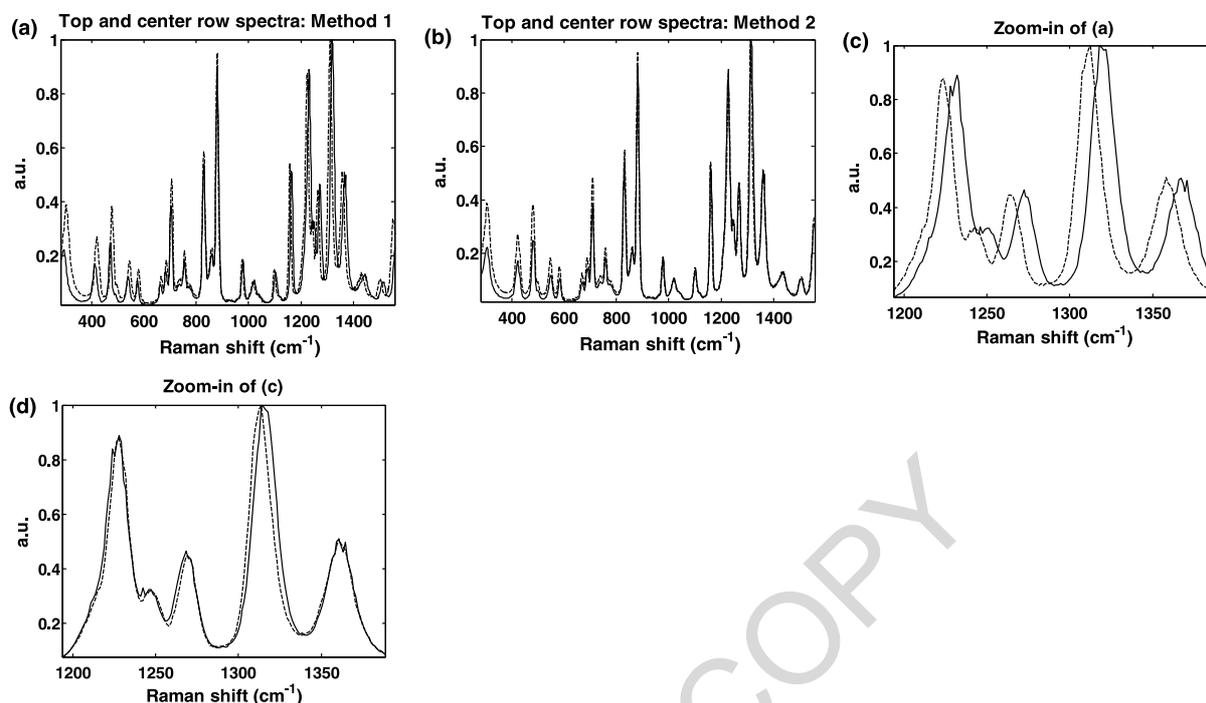
Fig. 4. Comparison of two spectra from the top ("–") and the center ("- -") row of the CCD: (a) After applying Method 1; (b) after applying Method 2; (c) zoom-in of high wavenumber region of (a); (d) zoom-in of high wavenumber region of (b).

binning to reduce the amount of data. The error introduced by the hardware binning is on average less than ~0.25 pixel, or equivalently, 2.5% of the spectral resolution (FWHM).

To further evaluate the improvement of Method 2 over Method 1, Fig. 4(a) shows the center-row spectrum and that of the top row after the application of Method 1. The error left uncorrected by Method 1 show up as apparent peak drifts. Because the curvature of the center wavelength was used for correction, the error becomes more significant towards either side of the CCD. Spectra from the same two rows are plotted after the application of Method 2 in Fig. 4(b), showing that the apparent peak drifts are greatly reduced, as better visualized in Fig. 4(c) and (d) for the high wavenumber region. The two spectra in Fig. 4(b) differ mainly in their intensity levels due to vignetting. As described above, system modifications simply require re-calibration to obtain the map and the scale factor, a great gain in flexibility.

Both Method 1 and Method 2 have been implemented for the home-built line-scan Raman microscope using acetaminophen powder as the calibration standard. Since the curvature is not as obvious as in the previous system due to the lower NA, lower dispersion, and smaller CCD size, both methods seem to perform well via visual inspection similar to those presented earlier, suggesting that the difference, if any, is not as apparent at the current CCD pixel resolution (Fig. 5). However, sub-pixel peak misalignments across the CCD have been detected after enhancing the original spectrum to 0.01 pixel resolution via interpolation. Figure 6 shows the mean and standard deviation of detected peak shifts in all CCD rows (1–198, 200–400) compared to the center row (199) of the CCD. We observe that Method 1 leaves sub-pixel (~−1 to 1) peak shifts un-corrected in 14 major peaks of acetaminophen except the one used for obtaining the curvature during the calibration step. This level of un-corrected peak shifts (about 10% of spectral resolution) is significant in applications where accurate quantitative information is needed, for example, multivariate calibration and spectral pattern recognition [3,4,6,18–20], and phonon scattering
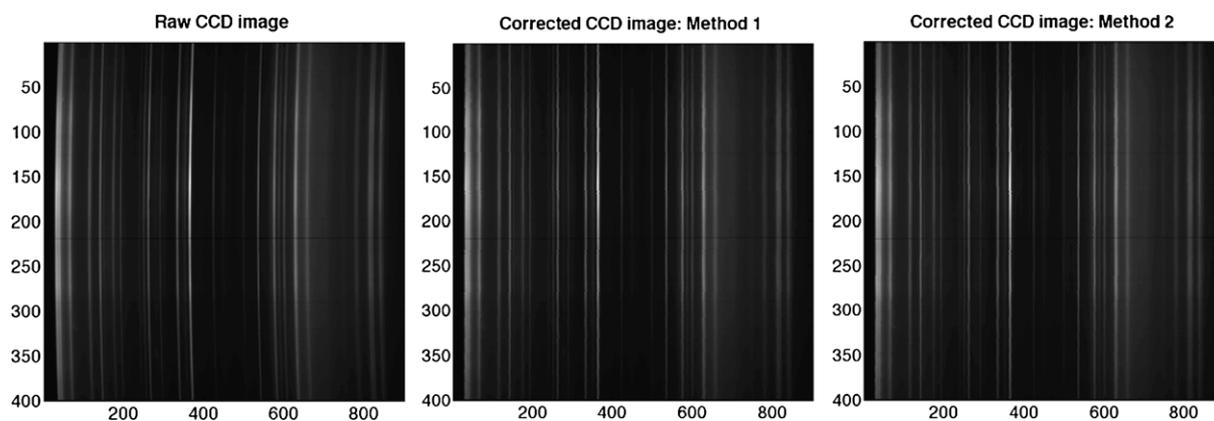
Fig. 5. CCD image of acetaminophen powder acquired by the line-scan Raman microscope: (a) raw image; (b) after applying Method 1; (c) after applying Method 2.
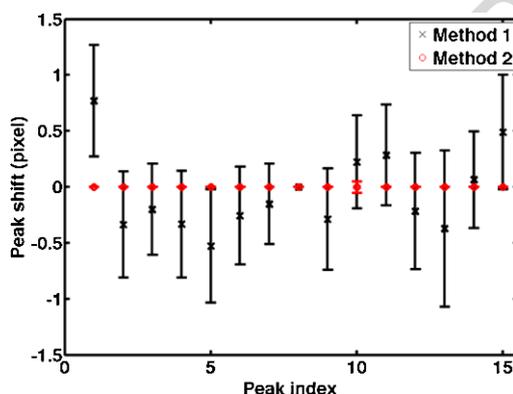
Fig. 6. Detected shifts in 15 major peaks of acetaminophen at sub-pixel resolution. (Colors are visible in the online version of the article; http://dx.doi.org/10.3233/BSI-140092.)

measurements based on peak shift for low-dimensional nanomaterials [1,2]. In contrast, Method 2 aligns all peaks nearly perfectly.

Direct comparison of peak misalignment and broadening has been employed in all the previous analysis. Although significant improvement has been observed, direct benefits for subsequent multivariate analysis have been uncertain. We therefore employ principal component analysis (PCA), a heavily used technique in the field, to assess the impact of curvature correction. We compare results using the previous dataset, as well as three other datasets acquired from disposed human adipose tissue, microparticles (3 μm polystyrene, Sigma-Aldrich), and a gold nanodot decorated surface-enhanced Raman scattering (SERS) substrate [10] coated with benzenethiol monolayers following protocols in [13]. These samples are representative for existing Raman and SERS applications. All data sets consist of 133 spectra obtained from our line-scan Raman microscope. Table 1 summarizes the total variance explained by the first 6 PCs in all data sets, suggesting that the uncorrected curvature can significantly distort variance distribution among principal components. By plotting the PC1 spectra from each data set, we can visualize the striking effect of curvature correction in Fig. 7. The spectra of the corrected data set appear to be much more representative for the molecular content of the sample. In particular, all peak splitting

Table 1

Percentage total variance explained by the first 6 principal components

	Acetaminophen		Adipose		Polystyrene		Benzenethiol	
	Raw	Corrected	Raw	Corrected	Raw	Corrected	Raw	Corrected
PC1	77.53	91.30	88.35	91.26	61.64	74.79	98.94	99.35
PC2	13.17	2.99	3.26	2.89	27.22	21.08	0.6	0.32
PC3	5.48	1.98	1.96	0.47	7.75	1.63	0.11	0.11
PC4	1.53	1.29	0.49	0.16	0.82	0.63	0.10	0.04
PC5	1.18	0.65	0.18	0.12	0.64	0.23	0.02	0.01
PC6	0.38	0.36	0.12	0.09	0.35	0.17	0.01	0.01
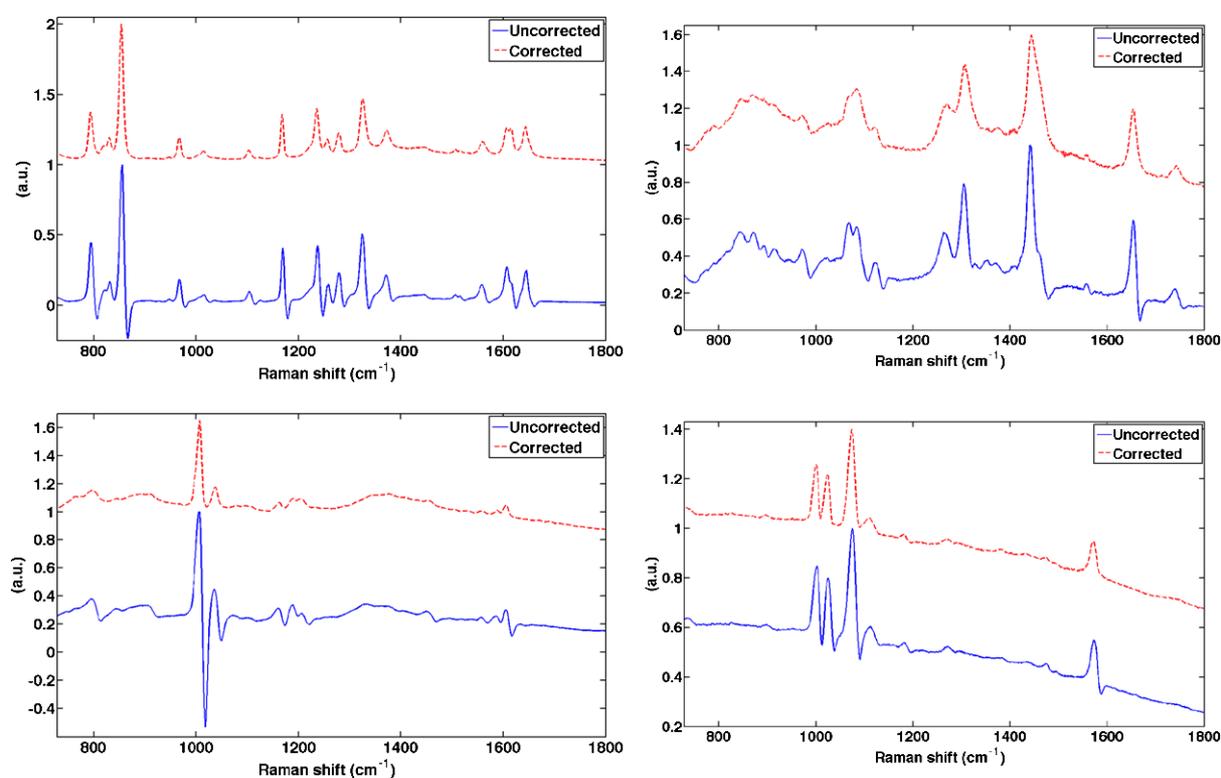
Fig. 7. Comparison of the first principal components (PC1) extracted from uncorrected and corrected data sets acquired from various types of samples: acetaminophen powder, adipose tissue, polystyrene microspheres, and benzenethiol self-assembled monolayer. (Colors are visible in the online version of the article; http://dx.doi.org/10.3233/BSI-140092.)

into negative values have been corrected. The results show that it is effective to use PC spectra directly to evaluate curvature correction methods.

An important issue during implementation is how accurately the calibration algorithm identifies the peaks serving as separation boundaries. We simulated scenarios for different amounts of random noise and found that peaks with sharp and well-defined lineshape and higher SNRs are more resistant to noise distortion, i.e., the true peak positions could be more accurately identified. Therefore, for practical implementation, the reference material must possess multiple distinctive peaks across the wavelength range of interest and the reference image for calibration must have superior SNR. Since the map for

correction is generated entirely based on the calibration data, the SNR in the application data has no influence on the performance.

4. Conclusion

For high-throughput Raman spectroscopy and microscopy, an elongated slit is generally employed together with a high numerical aperture (NA) grating spectrograph and large-format 2-dimensional charge-coupled device (CCD) camera. Due to out-of-plane light incidence at the grating, the linear slit image becomes approximately a parabola. Simply binning vertical pixels worsens spectral resolution and signal-to-noise ratio (SNR), and precautions must be taken. In this paper, we modeled and simulated the formation of the image curvature with actual parameters of our instrument. We compared the performance of two software-based curvature correction methods and showed improvement in experimental data fidelity in terms of spectral resolution and peak alignment on both spectroscopy and microscopy Raman systems. We further experimentally showed that by calibrating on multiple Raman peaks, better performance has been achieved on both systems. Implementation of this new method is through a two-step process: Image curvature calibration and application. Once the calibration is accomplished, the correction can be employed in future experiments, providing great flexibility for system modifications, i.e., only re-calibration is required.

Acknowledgements

References

[1] K.A. Alim, V.A. Fonoberov, M. Shamsa and A.A. Balandin, Micro-Raman investigation of optical phonons in ZnO nanocrystals, *Journal of Applied Physics* **97** (2005), 124313.

[2] A.A. Balandin, Thermal properties of graphene and nanostructured carbon materials, *Nature Materials* **10** (2011), 569–581.

[3] K.L. Bechtel, W.C. Shih and M.S. Feld, Intrinsic Raman spectroscopy for quantitative biological spectroscopy, Part II: Experimental applications, *Optics Express* **16** (2008), 12737–12745.

[4] A.M.K. Enejder, T.G. Scecina, J. Oh, M. Hunter, W.C. Shih, S. Sasic, G.L. Horowitz and M.S. Feld, Raman spectroscopy for noninvasive glucose measurements, *Journal of Biomedical Optics* **10** (2005), 031114.

[5] F.W.L. Esmonde-White, K.A. Esmonde-White and M.D. Morris, Minor distortions with major consequences: correcting distortions in imaging spectrographs, *Applied Spectroscopy* **65** (2011), 85–98.

[6] A.S. Haka, K.E. Shafer-Peltier, M. Fitzmaurice, J. Crowe, R.R. Dasari and M.S. Feld, Diagnosing breast cancer by using Raman spectroscopy, *Proceedings of the National Academy of Sciences of the United States of America* **102** (2005), 12371–12376.

[7] Z.W. Huang, H.S. Zeng, I. Hamzavi, D.I. McLean and H. Lui, Rapid near-infrared Raman spectroscopy system for real-time *in vivo* skin measurements, *Optics Letters* **26** (2001), 1782–1784.

[8] T.M. James, M. Schlosser, R.J. Lewis, S. Fischer, B. Bornschein and H.H. Telle, Automated quantitative spectroscopic analysis combining background subtraction, cosmic ray removal, and peak fitting, *Applied Spectroscopy* **67** (2013), 949–959.

[9] C.A. Lieber and A. Mahadevan-Jansen, Automated method for subtraction of fluorescence from biological Raman spectra, *Applied Spectroscopy* **57** (2003), 1363–1367.

[10] A. Merlen, V. Gadenne, J. Romann, V. Chevallier, L. Patrone and J.C. Valmalette, Surface enhanced Raman spectroscopy of organic molecules deposited on gold sputtered substrates, *Nanotechnology* **20** (2009), 215705.

[11] I. Pelletier, C. Pellerin, D.B. Chase and J.F. Rabolt, New developments in planar array infrared spectroscopy, *Applied Spectroscopy* **59** (2005), 156–163.

[12] J. Qi, J. Li and W.-C. Shih, High-speed hyperspectral Raman imaging for label-free compositional microanalysis, *Biomedical Optics Express* **4** (2013), 2376–2382.

[13] J. Qi, P. Motwani, M. Gheewala, C. Brennan, J.C. Wolfe and W.-C. Shih, Surface-enhanced Raman spectroscopy with monolithic nanoporous gold disk substrates, *Nanoscale* **5** (2013), 4105–4109.

[14] J. Qi, P. Motwani, J. Wolfe and W. Shih, High-throughput Raman and surface-enhanced Raman microscopy, in: *Proc. SPIE*, Vol. 8219, 2012, 821903.

[15] J. Qi and W.C. Shih, Parallel Raman microspectroscopy using programmable multipoint illumination, *Optics Letters* **37** (2012), 1289–1291.

[16] J. Qi and W.C. Shih, Performance of line-scan Raman microscopy (LSRM) for high-throughput chemical imaging of cell population, *Applied Optics* **53**(13) (2014), 2881–2885.

[17] H.G. Schulze and R.F.B. Turner, A fast automated polynomial-based cosmic ray spike-removal method for the high-throughput processing of Raman spectra, *Applied Spectroscopy* **67** (2013), 457–462.

[18] W.C. Shih, K.L. Bechtel and M.S. Feld, Constrained regularization: hybrid method for multivariate calibration, *Analytical Chemistry* **79** (2007), 234–239.

[19] W.C. Shih, K.L. Bechtel and M.S. Feld, Intrinsic Raman spectroscopy for quantitative biological spectroscopy, Part I: Theory and simulations, *Optics Express* **16** (2008), 12726–12736.

[20] W.C. Shih, K.L. Bechtel and M.S. Feld, *Handbook of Optical Sensing of Glucose in Biological Fluids and Tissues*, 2008.

[21] W.-C. Shih, K.L. Bechtel and M.S. Feld, Noninvasive glucose sensing with Raman spectroscopy, in: *Analytical Chemistry of in vivo Glucose Measurements*, Wiley, Hoboken, NJ, 2009, pp. 391–419.

[22] J. Qi, J. Zeng, F. Zhao, S.H. Lin, B. Raja, U. Strych, R.C. Willson and W.-C. Shih, Label-free, *in situ* SERS monitoring of individual DNA hybridization in microfluidics, *Nanoscale* **6**(15) (2014), 8521–8526.

[23] J. Zhao, Image curvature correction and cosmic removal for high-throughput dispersive Raman spectroscopy, *Applied Spectroscopy* **57** (2003), 1368–1375.