Sensors & Actuators: B. Chemical 299 (2019) 126902

Contents lists available at ScienceDirect

Sensors and Actuators B: Chemical

journal homepage: www.elsevier.com/locate/snb

Quaternion-based parallel feature extraction: Extending the horizon of quantitative analysis using TLC-SERS sensing

Yong Zhao[a,b], Ailing Tan[a,c], Kenny Squire[a], Kundan Sivashanmugan[a], Alan X. Wang[a,*]

[a] School of Electrical Engineering and Computer Science, Oregon State University, Corvallis, OR, 97331, USA
[b] School of Electrical Engineering, The Key Laboratory of Measurement Technology and Instrumentation of Hebei Province, Yanshan University, Qinhuangdao, Hebei, 066004, PR China
[c] School of Information Science and Engineering, The Key Laboratory for Special Fiber and Fiber Sensor of Hebei Province, Yanshan University, Qinhuangdao, Hebei, 066004, PR China

ARTICLE INFO

Keywords:
Thin layer chromatography
Surface-enhanced Raman spectroscopy
Quantitative analysis
Quaternion principal component analysis

ABSTRACT

Quantitative analysis using thin-layer chromatography coupled in tandem with surface-enhanced Raman scattering (TLC-SERS) still remains a grand challenge due to many uncontrollable variations during the TLC developing process and the random nature of the SERS substrates. Traditional chemometric methods solve this problem by sampling multiple SERS spectra in the sensing spot and then conducting statistical analysis of the SERS signals to mitigate the variation of quantitative analysis, while still ignoring the spatial distribution of the target species and the correlation among the multiple sampling points. In this paper, we proposed for the first time a parallel feature extraction and fusion method based on quaternion signal processing techniques, which can enable quantitative analysis using recently established TLC-SERS techniques. By marking three deterministic sampling points, we recorded spatially correlated SERS spectra to constitute an integral representation model of triple-spectra by a pure quaternion matrix. Quaternion principal component analysis (QPCA) was utilized for features extraction and followed by feature crossing among the quaternion principal components to obtain final fusion spectral feature vectors. Support vector regression (SVR) was then used to establish the quantitative model of melamine-contaminated milk samples with seven concentrations (1 ppm–250 ppm). Compared with traditional TLC-SERS analysis methods, QPCA method significantly improved the accuracy of quantification by reaching only 7% and 2% quantization errors at 20 and 105 ppm concentration. Validation testing based on reasonable amount of statistic measurement results showed consistently smaller measurement errors and variance, which proved the effectiveness of QPCA method for TLC-SERS based quantitative sensing applications.

1. Introduction

Thin layer chromatography combined with surface-enhanced Raman spectroscopy (TLC-SERS) has become a very effective technique for detecting target molecules from a mixture or a complex sample with exclusive advantages of simplicity, high throughput, and cost effectiveness [1]. Briefly, a mixture sample is spotted onto a commercial silica-gel TLC plate (stationary phase) or a specially designed porous substrate. The eluent (mobile phase) migrates through the TLC plate via capillary flow. Different molecules carried by the eluent flow will be separated due to different affinities toward the stationary and mobile phases. After TLC separation, gold or silver colloidal nanoparticles are casted or sprayed onto the concentrated spots and then SERS spectra will be collected by a Raman microscope or a Portable Raman spectrometer. In the past years, many TLC-SERS chemical sensing results have been reported including detecting artist dyes in fibers [2], substituted aromatic pollutants in water [3], apomorphine in human plasma [4], tobacco-related biomarkers and cocaine in urine [5,6], organophosphate pesticide in tea leaves [7], adulterants in botanical dietary supplements [8,9], pericarpium papaver in hot pot [10], aflatoxins in peanuts [11], Sudan-I in cooking oil [12], and chemical reaction [13].

Despite these enormous progress, accurate quantitative analysis using TLC-SERS technique is still a grand challenge due to two intrinsic natures associated with the TLC-SERS processes. First, the distribution of target molecules within each analyte concentration spot is non-uniform and can be disturbed by many uncontrollable factors such as the inhomogeneous porosity of the TLC plate, temperature and airflow fluctuation during the TLC process, and even the dispensing of the plasmonic colloids. Second, the measured SERS spectra from the

* Corresponding author.
E-mail address: wang@eecs.oregonstate.edu (A.X. Wang).

https://doi.org/10.1016/j.snb.2019.126902
Received 30 May 2019; Received in revised form 24 July 2019; Accepted 29 July 2019
Available online 03 August 2019
0925-4005/

deposited plasmonic nanoparticles are naturally random due to the hot spot strength variation, non-uniform adsorption of the molecules, and random scattering of the SERS signals from the porous TLC plate [14]. The porous TLC plate is also detrimental to high sensitivity analysis because only a portion of separated target molecules can be adsorbed by the plasmonic nanoparticles at the top surface of the TLC plate, which will contribute to the measurable SERS signals. To minimize the impact of the porous TLC plate, some researchers transferred the separated analyte onto a uniform SERS substrate by dissolving it with a suitable solvent [15]. Another group used thin layers of metal nanoparticles as both the TLC layer and the SERS substrate [16,17]. However, the transfer of the analyte spot from the TLC layer requires extra efforts and the usage of the metallic TLC plates could comprise the separation capability of the TLC plate. In addition to the optimization of TLC plates and SERS substrates, an internal reference method was also adopted to reduce the fluctuations of SERS signals [18]. Nevertheless, it still relies on single point spectrum to obtain the semi-quantitative model. Other groups proposed quantitative analysis based on digital images of the analyte spots in the TLC plate [19,20]. Although with potential to calibrate the TLC plate variation in theory, as an indirect measurement method, it cannot measure the spectra of target molecules and lack specificity of SERS sensing. In summary, existing quantitative TLC-SERS analysis methods are exclusively based single-point sampling or random multiple-point sampling. In reality, after the TLC separation, the diameter of the analyte concentration spot is usually much larger than that of the laser beam, which will result in spatial distribution of the SERS spectra. Ignoring the correlation between the SERS spectra and the spatial distribution of the analyte will lead to poor reproducibility and inaccurate quantitative analysis.

The quaternion is a member of noncommutative division algebra which was invented by William Rowan Hamilton [21]. Many quaternion-based algorithms have been proposed in the field of signal processing, including quaternion Fourier transform [22], quaternion singular value decomposition [23], quaternion principal component analysis [24,25], quaternion independent component analysis [26,27], quaternion singular spectrum analysis [28], and quaternion principal component analysis network [29]. These proposed algorithms have been widely applied in computer graphics [30], aerospace applications [31], image processing [32], pattern recognition [33,34], vector sensor signal processing [35], blind extraction [36], adaptive filtering [37], biomedical signals analysis [38], and hyperspectral imaging [39]. In this work, we introduce a quaternion-based parallel feature extraction method for multi-spectral processing of TLC-SERS analysis, which significantly improved the accuracy of quantitative analysis. As shown in Scheme 1, we prepared melamine contaminated milk samples with the concentrations ranging from 1 to 250 ppm and performed TLC separation. After that, we recorded the SERS spectra at the top edge, the center and the bottom edge positions inside the analyte spot using a portable Raman spectrometer. Next, a parallel representation model of the triple-spectral data was constructed using a pure quaternion matrix. Quaternion principal component analysis (QPCA) was utilized for feature extraction and followed by feature crossing [40] between the quaternion principal components to obtain final fusion of spectral feature vectors. Finally, a quantitative model was built using support vector regression (SVR) algorithm.

2. Materials and methods

2.1. Synthesis of gold nanoparticles

The gold nanoparticles (Au NPs) were synthesized by sodium citrate as the reducing and stabilizing agent developed by Grabar, K. C. [41]. Briefly, a total of 100 mL 1 mM chloroauric acid aqueous solution was heated to the boiling point under vigorous stirring. After adding 4.2 mL of 1% sodium citrate, the pale yellow solution turned fuchsia quickly. The colloids were kept under reflux for another 20 min to ensure complete reduction of Au ions. After cooling to room temperature, the colloids were centrifuged at 9000 rpm for 15 min. The UV–vis absorption spectra and the SEM image of the prepared Au NPs were shown in Figs. S1 and S2 respectively. From the UV–vis spectroscopy measurement, the localized surface plasmonic resonance peak is at 528 nm with a relatively narrow spectral width, which indicates their diameters of approximately 40 nm. According to the basis of the Lambert's law, the concentration of Au NPs was calculated to be 4×10^{-10} M with a molar extinction coefficient of $3.4 \times 10^{10}\ M^{-1}\ cm^{-1}$.

2.2. Samples preparation

Melamine solution with concentration of 500 ppm was prepared by dissolving the melamine powder in methanol and lower concentration solution was prepared via stepwise diluting with methanol. Fresh whole milk was purchased from local supermarket. Melamine-spiked milk samples were prepared by adding equal amounts (5 ml) of melamine solution and milk under vigorous stirring at room temperature to obtain homogeneous solutions with the final melamine concentration of 250, 100, 25, 10, 5, 2 and 1 ppm.

2.3. TLC-SERS experiment

TLC-SERS was performed in the following three steps. First, 2 μL sample solution was spotted by a micro-pipette at 12 mm from the edge of the TLC plate, which is Silica-Gel 60 from Merck Inc. The TLC plates used herein has very low fluorescence background under 380 nm UV light illumination. Drying naturally in the air, the plate was kept in a TLC development chamber and eluted by Acetone: Chloroform: Ammonia (7:0.5:2) mobile phase for 15 min. After the eluent on the TLC plate evaporated naturally, 2 μL Au NPs were drop casted onto the analyte spot, which is determined according to previously measured retention factor (R_f) as discussed in Section 3.1. Finally, a portable Raman spectrometer equipped with a diode laser emitting at 785 nm wavelength for illumination over a 100 μm diameter was used to obtain the SERS signals. Due to the low photon energy at 785 nm wavelength, the Raman laser minimizes the excitation of fluorescence background, which yields very clear SERS spectra. The laser power, the scanning range, the optical resolution, and the integration time were set as 30 mW, 400–1800 cm^{-1}, 2 cm^{-1} and 5000 ms respectively. After TLC separation, the diameter of the analyte spot is about 3 mm. In order to collect three SERS spectra within each analyte spot, we conducted sequential SERS measurement at the center of the analyte spot and ± 1.35 mm along the mobile phase direction using a positioning stage, which is the accessory of the i-Raman Plus Portable Fiber Optic Raman System. Each SERS measurement was repeated three times.

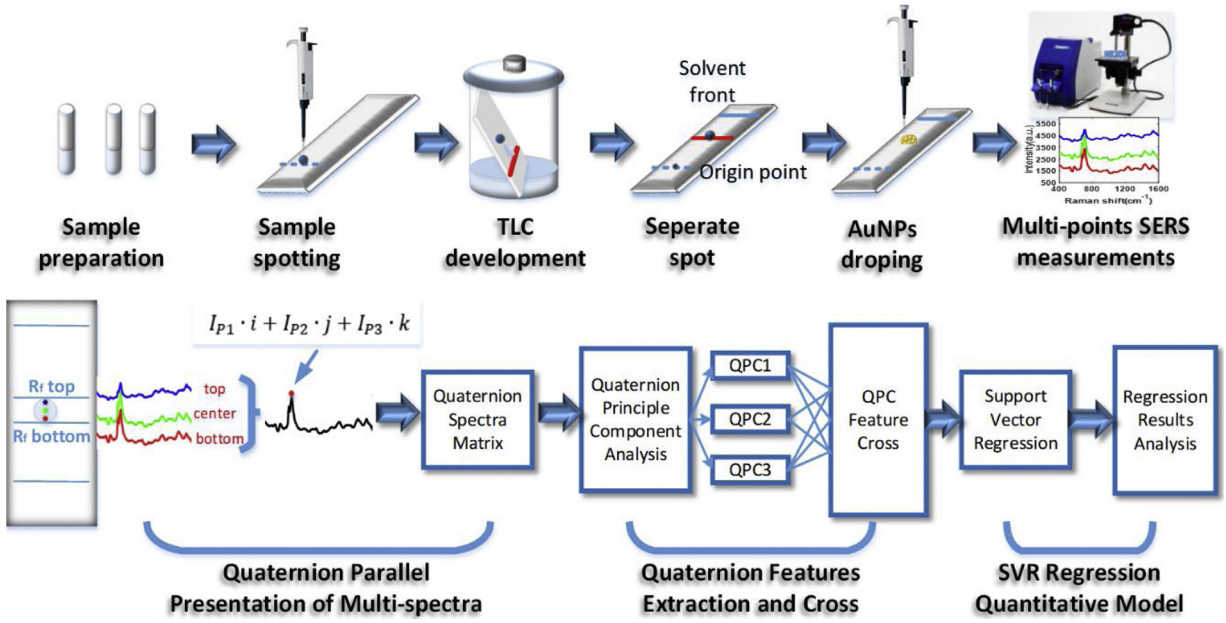
2.4. Multispectral quaternion parallel representation model

The quaternion toolbox for MATLAB is used in the TLC-SERS spectra analysis [42]. In order to process the spectral data of the three positions inside the analyte spot range simultaneously and maintain the correlation among them, we proposed a parallel triple-spectra representation model based on quaternion valued signals. The basic idea is utilizing the multi-dimensional quaternion domain and treat the triple-spectra as a single hyper-complex number. More introduction about the quaternion model are included in the S1 of algorithm explanation in the supporting information.

Here we encoded three SERS spectra intensity values at a specific wave number into a pure quaternion:

$$q(w_i) = I(w_i, p1)\bullet i + I(w_i, p2)\bullet j + I(w_i, p3)\bullet k \quad (1)$$

Where, $q(w_i)$ represents the triple-spectra as a complex number. Each spectral intensity $I(w, p1)$, $I(w, p2)$ and $I(w, p3)$ at three separated sequential measurement positions $p1(R_f\ bottom)$, $p2(R_f\ center)$ and $p3(R_f\ top)$ at the specific wavelength value w_i corresponds to the three



Scheme 1. TLC-SERS sensing of melamine-contaminated milk and quantitative analysis based on QPCA and feature crossing combined with SVR algorithm.

imaginary parts of the pure quaternion respectively.

In case of one sample, spectra of three measurement points throughout the entire wavelength range can be described as a pure quaternion vector as follows:

$$Q_t = [q_{w_1}, \cdots, q_{w_i}, \cdots q_{w_M}] \tag{2}$$

Where, w_i is the wavelength range of the SERS spectra, t = 1,2,⋯,N is the number of samples.

Considering a set of N samples of a specific concentration in the TLC-SERS experiment, all spectral data can be constructed as a quaternionic spectra matrix:

$$Q = \begin{bmatrix} Q_1^T \\ \vdots \\ Q_t^T \\ \vdots \\ Q_N^T \end{bmatrix} \tag{3}$$

Where, $Q \in \mathbb{Q}^{N \times M}$, is a matrix with the size of $N \times M$ in which each element is a pure quaternion, M is the number of the wavelength.

Based on this quaternion multi-spectral model, we can analyze TLC-SERS spectral data using quaternion signal processing techniques over the quaternion domainQ. At the same time, this model can preserve correlation among the triple-points of the TLC process with the advantage of quaternion domain for representing three dimensional signals in a natural way.

3. Results and discussion

3.1. TLC-SERS spectra of melamine in real milk samples

After separation of melamine-spiked milk samples with seven different concentrations, the melamine spot was visualized by iodine colorimetry and the range of the R_f was measured to be 0.57–0.63 and shown in Fig. 1(b). Since the Rf values are quite deterministic, it does not require further iodine colorimetry for real sample measurement. We record the SERS spectra at the top edge, center and bottom edge within the spot range respectively. For the spiked milk sample with the melamine concentration of 250 ppm, the triple-points SERS spectra inside the analyte spot were shown in Fig. 1(a). The Raman peak at 709.8 cm^{-1} of melamine could be clearly seen in the SERS spectra. Triple-points TLC-SERS spectra of other concentrations could be found in Fig. S3 of the Supporting Information.

3.2. QPCA feature extraction and feature cross

To the quaternion spectra matrix, QPCA was first conducted. More introduction about QPCA could be found in the S2. After quaternion principal component decomposition of raw quaternion spectral matrix, we obtain the first three quaternion principal component features QPC_1, QPC_2 and QPC_3, which are quaternion forms as follows:

$$QPC_i = P_{i1} + P_{i2} \bullet i + P_{i3} \bullet j + P_{i4} \bullet k \ (i = 1,2,3) \tag{4}$$

After normalization of the quaternion principal components and transformation from quaternion to Euler angle, we can visualize them onto the surface of a unit sphere, which were shown in Fig. 2. The details about the visualization method are included in the S2 of algorithm explanation in the supporting information.

From Fig. 2, it is still difficult to distinguish different concentrations based on single quaternion principal component feature. Therefore, we performed feature cross and fusion operation with the first three quaternion principal components. QPC_1, QPC_2, and QPC_3. The operator that obtains the scalar part of a quaternion is abbreviated as $\langle \bullet \rangle_R$ and that obtain the three imaginary parts of a quaternion is abbreviated as $\langle \bullet \rangle_I$. Then the first-order feature vector was constructed with three fusion modes, which extracted the real parts, imaginary parts and both parts of each quaternion principal component respectively.

$$Fea_1^{1ord} = [\langle QPC_1 \rangle_R, \langle QPC_2 \rangle_R, \langle QPC_3 \rangle_R] \in R^3 \tag{5}$$

$$Fea_2^{1ord} = [\langle QPC_1 \rangle_I, \langle QPC_2 \rangle_I, \langle QPC_3 \rangle_I] \in R^3 \tag{6}$$

$$Fea_3^{1ord} = [\langle QPC_1 \rangle_R, \langle QPC_1 \rangle_I, \langle QPC_2 \rangle_R, \langle QPC_2 \rangle_I, \langle QPC_3 \rangle_R, \langle QPC_3 \rangle_I] \in R^6 \tag{7}$$

First, the second-order feature vector was constructed with feature cross operation. A feature cross is a synthetic feature formed by multiplying two or more features. In this paper we multiply two different quaternion principal components extracted from first three quaternion principal components in rotation. Then we got the real and imaginary part using the operators defined by $\langle \bullet \rangle_R$ and $\langle \bullet \rangle_I$ operators for each synthetic crossing feature to get the final feature vector. This procedure was illustrated as follows:

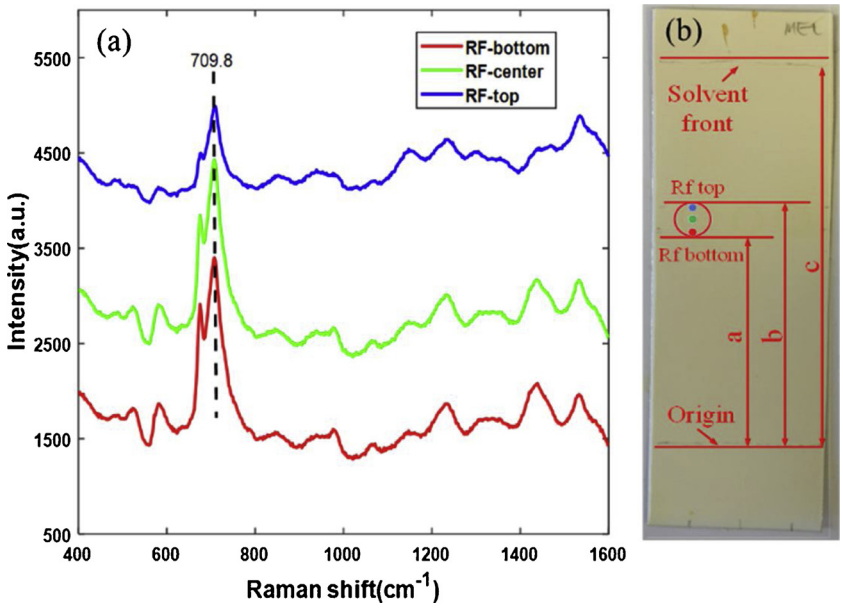


Fig. 1. (a) TLC-SERS spectra of melamine spiked milk (250 ppm) at different measurement positions (b)TLC plate shown the measurement positions R_f bottom, center and top, respectively.

$$FeaCross_1 = QPC_1 \bullet QPC_2 \in \mathbb{Q} \tag{8}$$

$$FeaCross_2 = QPC_1 \bullet QPC_3 \in \mathbb{Q} \tag{9}$$

$$FeaCross_3 = QPC_2 \bullet QPC_3 \in \mathbb{Q} \tag{10}$$

$$Fea_1^{2ord} = [\langle FeaCross_1\rangle_R, \langle FeaCross_2\rangle_R, \langle FeaCross_3\rangle_R] \in R^3 \tag{11}$$

$$Fea_2^{2ord} = [\langle FeaCross_1\rangle_I, \langle FeaCross_2\rangle_I, \langle FeaCross_3\rangle_I] \in R^3 \tag{12}$$

$$Fea_3^{2ord} = [\langle FeaCross_1\rangle_R, \langle FeaCross_1\rangle_I, \langle FeaCross_2\rangle_R, \langle FeaCross_2\rangle_I, \langle FeaCross_3\rangle_R, \langle FeaCross_3\rangle_I] \in R^6 \tag{13}$$

We performed quantitative regression analysis using SVR with the aforementioned three feature-fusion modes respectively and obtained the quantitative analysis results that were shown in the Table1.

From Table1, we can conclude that the best regression performance can be obtained based on second-order synthetic features because cross features actually can represent the synergy of the respective information of a single quaternion principal component, which can provide better nonlinear expression abilities beyond which a single feature can achieve individually. The principal component with the real-part features of the first-order and second-order cross features were plotted respectively in the Fig. 3(a) and (b). From Fig. 3, it is obvious that the seven groups of spectra with different concentrations can be separated very clearly. At the same time, the similar regression effect was obtained based on feature vectors combined with the first-order and second-order real-part features, which integrated the whole feature information of raw spectra dataset.

Table 1
Quantitative analysis results with different feature-cross and fusion methods.

Feature cross and fusion	Training Set		Testing Set		
	RMSECV	R^2	RMSEP	R^2	RPD
Fea_1^{1ord}	3.1095	0.99921	8.0089	0.99004	12.842
Fea_2^{1ord}	34.306	0.81121	42.047	0.75264	2.0166
Fea_3^{1ord}	2.8646	0.9986	9.9674	0.98587	9.0359
Fea_1^{2ord}	0.72778	0.99994	3.9749	0.99762	24.246
Fea_2^{2ord}	30.986	0.87595	31.041	0.86554	2.4774
Fea_3^{2ord}	5.3651	0.99619	9.9660	0.98527	9.2044
$[Fea_1^{1ord}, Fea_1^{2ord}]$	0.19785	0.99988	4.1864	0.99722	23.792
$[Fea_2^{1ord}, Fea_2^{2ord}]$	28.903	0.8744	31.969	0.85347	2.6343
$[Fea_3^{1ord}, Fea_3^{2ord}]$	4.4452	0.99775	10.168	0.98549	8.4949

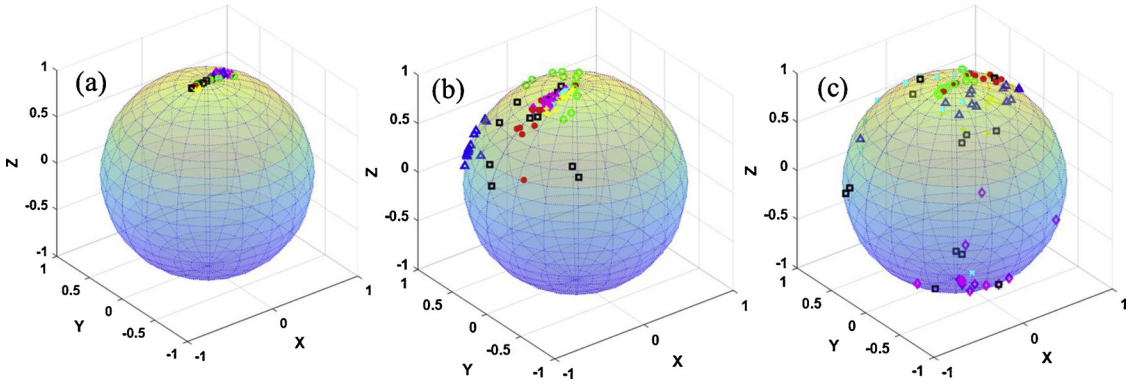


Fig. 2. Visualization of quaternion principal components on a unit sphere surface. (a), (b), (c) are the first, second and third quaternion principal component, respectively.

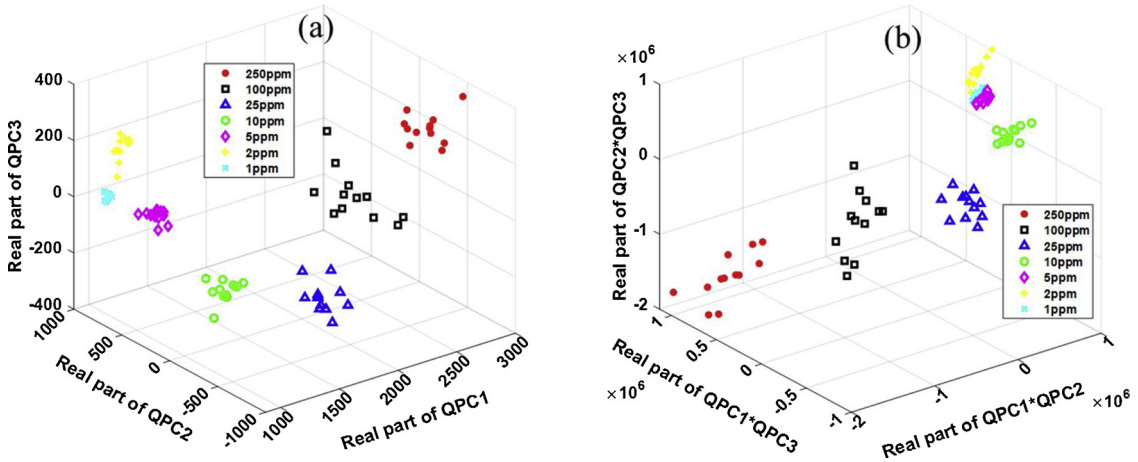


Fig. 3. Quaternion principal components scatter plot of the real-part features (a) the first-order quaternion principal components and (b) the second-order quaternion principal components.

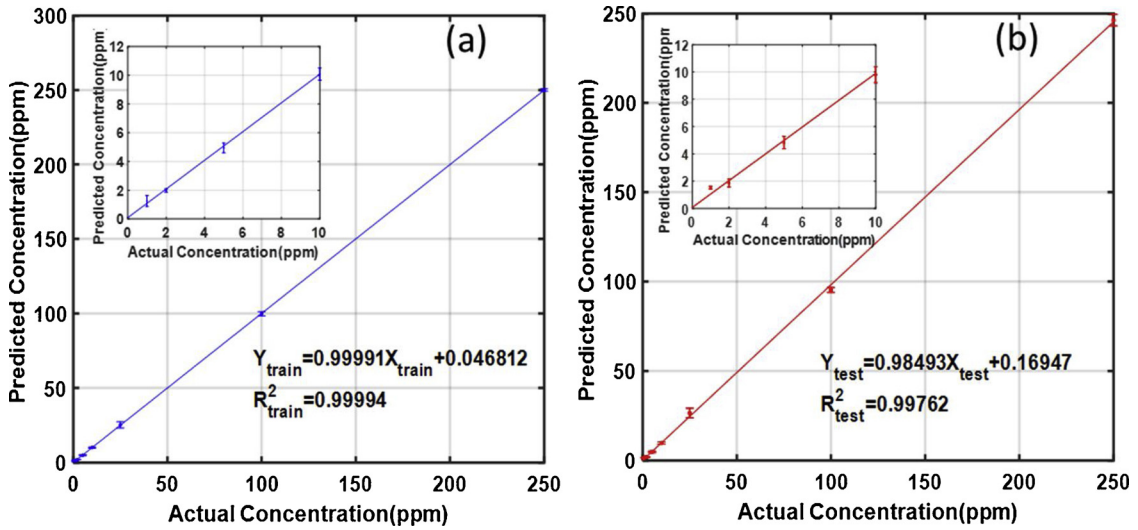


Fig. 4. Quantitative fitting of actual concentrations and predicted concentrations of melamine in milk using quaternion-based feature extraction combined with SVR model: (a) training dataset; and (b) testing dataset.

Table 2
Comparison between the serial processing method based on single channel spectrum and proposed parallel processing method based on quaternion feature extraction.

Process Method	Optimal PCs	Training Set		Testing Set		
		RMSECV	R^2	RMSEP	R^2	RPD
Channel1	4	7.3014	0.99108	10.113	0.98498	9.2667
Channel2	3	5.9936	0.98757	12.684	0.97603	7.3982
Channel3	12	12.039	0.98474	31.932	0.85404	2.5856
Serial Channels	19	0.15123	0.99997	9.6414	0.98656	10.033
Sum Channels	3	27.599	0.87375	31.905	0.85651	2.4111
Average Channels	10	5.1767	0.99426	11.866	0.98057	7.3578
Quaternion Channels	3	0.1955	0.99999	3.8131	0.9974	29.271

3.3. Multivariate Regression using support vector regression

We employed SVR method to obtain the regression prediction model for the melamine contaminated milk samples. As the nonlinear characteristics originated from the complex procedure of the TLC development and the variation of the SERS spectra, nonlinear regression method is suitable for TLC-SERS quantitative modeling. RBF kernel was chosen in constructing the SVR model. The grid searching method was conducted to determine the optimal values of two key parameters (γ for the RBF kernel and C for the SVM) in the searching range [2^{-10}–2^{10}]. The parameters pair with the best cross-validation accuracy is determined to be (12.1257, 1024). Based on the optimal parameters, the calibration curves for the actual concentrations and predicted concentrations with the training and testing data set were shown in Fig. 4(a) and (b) with the zoomed-in view plots of calibration curves in the low concentration range (1 ppm to 10 ppm). The mean and variance of the prediction results were shown in Table S1 of the Supporting Information. It can be seen that the predicted concentrations were very close to the actual concentrations for each sample for the entire concentration range of 1–250 ppm.

In order to prove the superiority of the proposed algorithm, we compared the results with those obtained by six commonly used chemometric methods based on three independent point spectral data. The first three comparison methods (Channel 1–3) utilized the SERS spectra from three uncorrelated measurement points. The Series Channels method concatenated the principal components of the three previous channels into a single long serial feature vector. The Summation Channels and Average Channels used the cumulative and average value of the three SERS spectra to obtain principal components respectively.

Table 3
Comparison of the validation results with mean and variance values.

Concentration	Predicted results	C_1	C_2	C_3	C_{serial}	C_{mean}	C_{sum}	$C_{quaternion}$
20ppm	Mean	18.9438	21.9972	39.0973	32.9074	34.1045	53.4921	21.4722
	Variance	5.2052	2.5925	3.4751	2.5857	5.1529	12.8846	2.3112
105ppm	Mean	93.0887	81.4893	93.4377	98.5279	81.6958	91.4590	101.8571
	Variance	4.8163	1.4430	7.0837	2.5064	0.9033	4.7776	0.4964

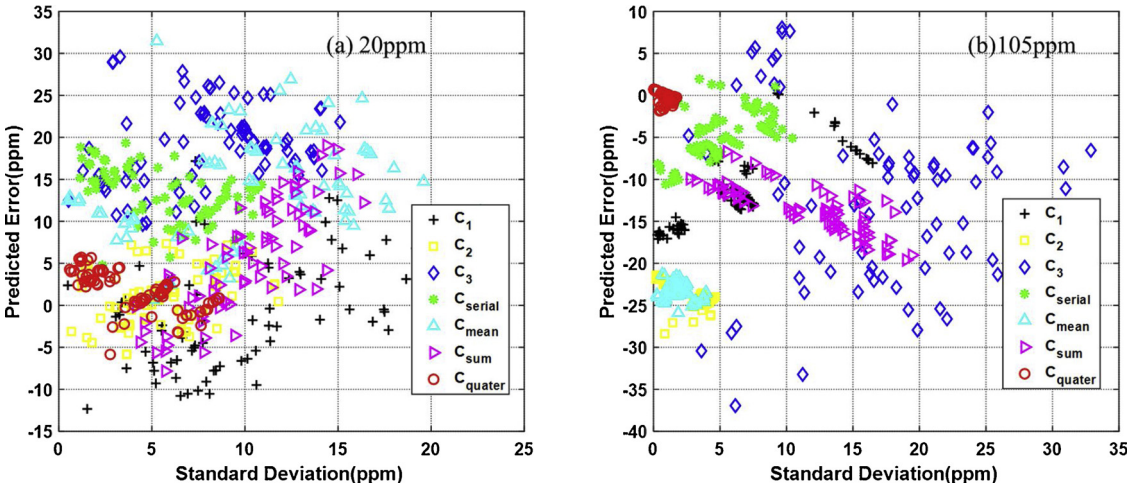


Fig. 5. Scattering plot of the validation results based on the models established with different methods: (a) 20 ppm and (b) 105 ppm.

Nevertheless, all these methods ignored the correlation of the SERS spectra and the spatial distribution. The comparison between the six aforementioned methods with the quaternion-based method was shown in Table2. Each comparison method was based on its respective optimal parameters and the relationship between regression performance and number of principal components was listed in Fig. S4 of the Supporting Information.

From analysis based on single-point measurement (Channel1–Channel3), we can see large variations of the quantitative fitting associated with the random nature of TLC-SERS sensing mechanism, which can constraint quantitative sensing for practical applications. The analysis based on statistical analysis, however, shows mixing outcome. Sum Channels using feature vector extracted from cumulative spectra led to even worse quantification. In contrast, the reliability and robustness will be better based on Average Channels and Serial Channels feature. Obviously, the quaternion-based parallel processing method can greatly improve the regression than the other six methods. The main reason is that the spatial distribution information of the TLC development was inherently encoded in the quaternionic representation model of multispectral data. The spectral data of three SERS sensing positions are treated jointly as a whole with their intra-correlations being fully taken into account. The quaternion feature extraction of multispectral data without losing the information between different detecting positions and the synthesized features obtained by feature crossing accurately reflect the nonlinear relationship embedded in multi-points spectral features of TLC development.

3.4. Validation with real samples

In order to verify the generalization ability of the model based on quaternion parallel feature extraction, 20 ppm and 105 ppm melamine-contaminated samples were prepared for blind testing. Without knowing the real melamine concentration, the second author Ailing Tan performed the TLC-SERS measurement according to the same procedure described in Section 2.3. Each sample were measured 15 times by TLC-SERS to evaluate the accuracy possibility. Next, the first author Yong Zhao processed the spectral data according to the methods described in this paper. Briefly, the quaternion representation spectral data is projected to the quaternion principal component projection matrix, and the quaternion principal component features were obtained. Then the features were crossed to obtain the fusion feature and the feature vector was used as input to the established SVR regression model. For each concentration group, 3 spectral data were selected randomly from the 15 spectra of each sample to predict the quantification results, effectively presenting 455 possible validation tests. The predicted mean values and standard variance values of the Quaternion Channels proposed in this paper and the comparative prediction results with the other methods were shown in Table 3. The measurement error between the mean value and the actual concentration is 7% for 20 ppm sample and only 2% for 105 ppm sample. It is very obvious that Quaternion Channels method achieved much more accurate mean values and the variance values were much smaller as well compared with other commonly used chemometric methods.

In order to show the statistic performance of the quantitative model established by the proposed method, the scattering plots of 80 randomly chosen predicted results were shown in Fig. 5(a) and (b) for 20 and 105 ppm, respectively. The horizontal axis is the standard deviation and the vertical axis is the difference of the mean value to the actual value of melamine concentration. As one can see from Fig. 5, all the predicted results based on quaternion parallel processing method are more closely distributed around the original point, representing better accuracy and smaller variance. The validation results can prove that the parallel quaternion principal component method is truly a powerful tool for quantitative TLC-SERS sensing.

4. Conclusions

This paper proposed for the first time a quaternion-based parallel feature extraction method that enables quantitative analysis using well-established TLC-SERS techniques, which suffer poor quantification accuracy. By collecting SERS spectral data from three deterministic points at each measurement, the multi-spectra were holistically expressed as a

pure quaternion matrix to preserve the spatial and coupling information of the TLC-SERS data. Followed by quaternion principal component feature extraction and feature cross, SVR regression method was utilized to obtain quantitative analysis results of melamine-contaminated milk with concentrations from 1 to 250 ppm. Compared with traditional TLC-SERS chemometric analysis methods, QPCA method significantly improved the accuracy of quantification in both the modeling test and validation test. Our research proved the effectiveness of QPCA method to enable TLC-SERS quantitative sensing, which will play critical roles in food safety, environmental protection, drug detection, homeland security, and forensics investigation.

Acknowledgements

The authors would like to acknowledge the support from the National Institutes of Health under Grant No. 1R21DA0437131, the Unites States Department of Agriculture under Grant No. 2017- 67021-26606 and the National Science Foundation under Grant No. 1701329. A. Tan and Y. Zhao would also like to acknowledge the support from China Scholarship Council.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.snb.2019.126902.

References

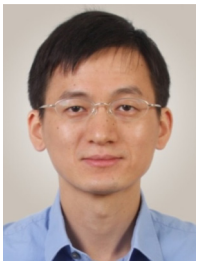
[1] W.W. Yu, I.M. White, Chromatographic separation and detection of target analytes from complex samples using inkjet printed SERS substrates, Analyst 138 (2013) 3679–3686.
[2] C.L. Brosseau, A. Gambardella, F. Casadio, C.M. Grzywacz, J. Wouters, R.P. Van Duyne, Ad-hoc surface-enhanced raman spectroscopy methodologies for the detection of artist dyestuffs: thin layer chromatography-surface enhanced raman spectroscopy and in situ on the Fiber analysis, Anal. Chem. 81 (2009) 3056–3062.
[3] D. Li, L. Qu, W. Zhai, J. Xue, J.S. Fossey, Y. Long, Facile on-site detection of substituted aromatic pollutants in water using thin layer chromatography combined with surface-enhanced raman spectroscopy, Environ. Sci. Technol. 45 (2011) 4046–4052.
[4] A. Lucotti, M. Tommasini, M. Casella, A. Morganti, F. Gramatica, G. Zerbi, TLC–surface enhanced Raman scattering of apomorphine in human plasma, Vib. Spectrosc. 62 (2012) 286–291.
[5] R. Huang, S. Han, X. Li, Detection of tobacco-related biomarkers in urine samples by surface-enhanced Raman spectroscopy coupled with thin-layer chromatography, Anal. Bioanal. Chem. 405 (2013) 6815–6822.
[6] X. Kong, X. Chong, K. Squire, A.X. Wang, Microfluidic diatomite analytical devices for illicit drug sensing with ppb-Level sensitivity, Sens. Actuators B Chem. 259 (2018) 587–595.
[7] C. Yao, F. Cheng, C. Wang, Y. Wang, X. Guo, Z. Gong, et al., Separation, identification and fast determination of organophosphate pesticide methidathion in tea leaves by thin layer chromatography–surface-enhanced Raman scattering, Anal. Methods 5 (2013) 5560–5564.
[8] D. Lv, Y. Cao, Z. Lou, S. Li, X. Chen, Y. Chai, et al., Rapid on-site detection of ephedrine and its analogues used as adulterants in slimming dietary supplements by TLC-SERS, Anal. Bioanal. Chem. 407 (2015) 1313–1325.
[9] Y. Liu, F. Lu, Adulterated pharmaceutical chemicals in botanical dietary supplements: novel screening approaches, Rev. Anal. Chem. 36 (3) (2017) 1–14.
[10] X. Hu, G. Fang, A. Han, J. Liu, S. Wang, Rapid detection of Pericarpium papaveris in hot pot condiments using thin-layer chromatography and surface enhanced Raman spectroscopy combined with a support vector machine, Anal. Methods 9 (2017) 2177–2182.
[11] L.-L. Qu, Q. Jia, C. Liu, W. Wang, L. Duan, G. Yang, et al., Thin layer chromatography combined with surface-enhanced raman spectroscopy for rapid sensing aflatoxins, J. Chromatogr. A 1579 (2018) 115–120.
[12] X. Kong, K. Squire, X. Chong, A.X. Wang, Ultra-sensitive lab-on-a-chip detection of Sudan I in food using plasmonics-enhanced diatomaceous thin film, Food Control 79 (2017) 258–265.
[13] Z.-M. Zhang, J.-F. Liu, R. Liu, J.-F. Sun, G.-H. Wei, Thin layer chromatography coupled with surface-enhanced raman scattering as a facile method for on-site quantitative monitoring of chemical reactions, Anal. Chem. 86 (2014) 7286–7292.
[14] Y. Jiao, J.D. Ryckman, P.N. Ciesielski, C.A. Escobar, G.K. Jennings, S.M. Weiss, Patterned nanoporous gold as an effective SERS template, Nanotechnology 22 (2011) 295302.
[15] C.E. Freye, N.A. Crane, T.B. Kirchner, M.J. Sepaniak, Surface enhanced raman scattering imaging of developed thin-layer chromatography plates, Anal. Chem. 85 (2013) 3991–3998.
[16] H. Takei, J. Saito, K. Kato, H. Vieker, A. Beyer, A. Gölzhäuser, et al., TLC-SERS plates with a built-in SERS layer consisting of cap-shaped noble metal nanoparticles intended for environmental monitoring and food safety assurance, . 4 (2015) 9.
[17] C. Wang, F. Cheng, Y. Wang, Z. Gong, M. Fan, J. Hu, Single point calibration for semi-quantitative screening based on an internal reference in thin layer chromatography-SERS: the case of Rhodamine B in chili oil, Anal. Methods 6 (2014) 7218–7223.
[18] D. Zhang, L. Huang, B. Liu, H. Ni, L. Sun, E. Su, et al., Quantitative and ultrasensitive detection of multiplex cardiac biomarkers in lateral flow assay with core-shell SERS nanotags, Biosens. Bioelectron. 106 (2018) 204–211.
[19] G.J. Van Berkel, V. Kertesz, Automated Sampling and Imaging of Analytes Separated on Thin-Layer Chromatography Plates Using Desorption Electrospray Ionization Mass Spectrometry, Anal. Chem. 78 (2006) 4938–4944.
[20] B. Hemmateenejad, N. Mobaraki, F. Shakerizadeh-Shirazi, R. Miri, Multivariate image analysis-thin layer chromatography (MIA-TLC) for simultaneous determination of co-eluting components, Analyst 135 (2010) 1747–1758.
[21] W.R. Hamilton, On a New species of imaginary quantities, connected with the theory of quaternions, Proceedings of the Royal Irish Academy 1836–1869 (2) (1840) 424–434.
[22] T.A. Ell, S.J. Sangwine, Hypercomplex Fourier Transforms of Color Images, Ieee Trans. Image Process. 16 (2007) 22–35.
[23] S.J. Sangwine, N. Le Bihan, Quaternion singular value decomposition based on bidiagonalization to a real or complex matrix using quaternion Householder transformations, Appl. Math. Comput. 182 (2006) 727–738.
[24] R. Zeng, J. Wu, Z. Shao, L. Senhadji, H. Shu, Quaternion softmax classifier, Electron. Lett. 50 (2014) 1929–1931.
[25] X. Xu, Z. Guo, C. Song, Y. Li, Multispectral palmprint recognition using a quaternion matrix, Sensors (Basel, Switzerland 12 (2012) 4633–4647.
[26] G.M. Menanno, N. Le Bihan, Quaternion polynomial matrix diagonalization for the separation of polarized convolutive mixture, Signal Processing 90 (2010) 2219–2231.
[27] J. Via, L. Vielva, I. Santamaria, D.P. Palomar, Independent component analysis of quaternion Gaussian vectors, 2010, IEEE Sens. Array Multichannel Signal Process. Workshop (2010) 145–148.
[28] B.B.D. Sacchi, Quaternion multichannel SSA for multicomponent seismic data, GeoConvention (2018) 1–5.
[29] R. Zeng, J. Wu, Z. Shao, Y. Chen, B. Chen, L. Senhadji, et al., Color image classification via quaternion principal component analysis network, Neurocomputing 216 (2016) 416–428.
[30] R. Mukundan, Quaternions: from classical mechanics to computer graphics, and beyond, Proceedings of the 7th Asian Technology Conference in Mathematics, (2002), pp. 97–98.
[31] L. Fortuna, G. Muscato, M.G. Xibilia, A comparison between HMLP and HRBF for attitude control, IEEE Trans. Neural Netw. 12 (2001) 318–328.
[32] P. Denis, P. Carre, C. Fernandez-Maloigne, Spatial and spectral quaternionic approaches for colour images, Comput. Vis. Image Underst. 107 (2007) 74–87.
[33] Z. Shao, H. Shu, J. Wu, B. Chen, J.L. Coatrieux, Quaternion Bessel–Fourier moments and their invariant descriptors for object reconstruction and recognition, Pattern Recognit. 47 (2014) 603–611.
[34] C. Singh, J. Singh, Multi-channel versus quaternion orthogonal rotation invariant moments for color image representation, Digit. Signal Process. 78 (2018) 376–392.
[35] S. Miron, N.L. Bihan, J.I. Mars, Quaternion-MUSIC for vector-sensor array processing, Ieee Trans. Signal Process. 54 (2006) 1218–1229.
[36] J. Via, D.P. Palomar, L. Vielva, I. Santamaria, Quaternion ICA From Second-Order Statistics, Ieee Trans. Signal Process. 59 (2011) 1586–1600.
[37] F. Ortolani, D. Comminiello, M. Scarpiniti, A. Uncini, Frequency domain quaternion adaptive filters: algorithms and convergence performance, Signal Processing 136 (2017) 69–80.
[38] S. Enshaeifar, S. Kouchaki, C.C. Took, S. Sanei, Quaternion Singular Spectrum Analysis of Electroencephalogram With Application in Sleep Analysis, Ieee Trans. Neural Syst. Rehabil. Eng. 24 (2016) 57–67.
[39] H. Li, H. Li, L. Zhang, Quaternion-based multiscale analysis for feature extraction of hyperspectral images, Ieee Trans. Signal Process. 67 (2019) 1418–1430.
[40] https://developers.google.com/machine-learning/crash-course/feature-crosses/video-lecture.
[41] K.C. Grabar, R.G. Freeman, M.B. Hommer, M.J. Natan, Preparation and characterization of Au colloid monolayers, Anal. Chem. 67 (1995) 735–743.
[42] S.J. Sangwine, Quaternion Toolbox for Matlab, (2015) http://qtfm.sourceforge.net/.

Yong Zhao is an Associate Professor of the School of Electrical Engineering at Yanshan University since 2012. He received his Ph.D. degree in Instrument Science and Technology from Yanshan University in 2012. From April 2018 to April 2019, he was a visiting scholar at the School of Electrical Engineering and Computer Science, Oregon State University. His current research interests focus on chemometrics, machine learning and deep learning for optical sensors and optical spectral data analysis.

Ailing Tan is an Associate Professor of the School of Information Science and Engineering at Yanshan University since 2013. She received her Ph.D. degree in Optical Engineering from Yanshan University in 2013. From April 2018 to April 2019, she was a visiting scholar at the School of Electrical Engineering and Computer Science, Oregon State University. Her research focuses on optical sensors and surface-enhanced Raman spectroscopy technology for chemical and biological detection.

Kenny Squire is a PhD candidate in the School of Electrical Engineering and Computer Science at Oregon State University. He received his B.S. degree in Electrical Engineering from Brigham Young University in 2015. His research focuses on optical biosensors utilizing surface plasmon resonance and surface-enhanced Raman spectroscopy.

Kundan Sivashanmugan is currently a Postdoctoral Research Fellow at Electrical Engineering and Computer Science at Oregon State University. He also worked at the Department of Chemical Engineering, National Cheng Kung University (NCKU), Taiwan as a Research Fellow from Dec. 2015-March 2018. He received his doctoral degree at the Department of Materials Science and Engineering, NCKU, Taiwan in 2015. His current research interest is focused on the fabrication of SERS-active substrates as a fast-screening detection platform for biological and toxic-chemical targets.



Alan X. Wang is an Associate Professor of the School of Electrical Engineering and Computer Science at Oregon State University since 2011. He received his Ph.D. degree in Electrical and Computer Engineering from the University of Texas at Austin in 2006. From 2007-2011, he was with Omega Optics, Inc., where he served as the Chief Research Scientist for 9 SBIR/STTR projects. His research interests include nanophotonic devices for optical interconnects, and optical sensors for chemical and biological detection. His current research activities are sponsored by the National Science Foundation, the National Institutes of Health, the Air Force Office of Scientific Research, the National Energy Technology Laboratory, and industrial sponsors such as Hewlett-Packard. His current research focuses on integrated and nanophotonic devices, optical biosensors for chemical and biological sensing, and free space optical communication system. He has more than 90 journal publications and 100 conference presentations, and also holds five U.S. patents. He is a senior member of IEEE Photonics, SPIE and OSA.