Toward-Automated Etalon Artefacts Elimination in Dual-Wavelength Raman Spectroscopy for In Vivo Applications

Wenyi Xu[a,b], Renzhe Bi[a], Yi Qi[a], Ruochong Zhang[a], Poongkulali Rajarahm[a], Alicia Ann May Yap[a], Dinish U.S[a], Qian Cheng[b*], Steven Thng[c*], Malini Olivo[a*]

[a] A*STAR Skin Research Labs (A*SRL), 31 Biopolis Way, Singapore 138669, Singapore.

[b] Institute of Acoustics, School of Physics Science and Engineering, Tongji University, Shanghai 200092, China.

[c] National Skin Center, Singapore 308205, Singapore.

[*] Email: malini_olivo@asrl.a-star.edu.sg

[*] Email: steventhng@nsc.com.sg

[*] Email: q.cheng@tongji.edu.cn

ABSTRACT: Confocal Raman spectroscopy (CRS) has emerged as a powerful tool for non-invasive chemical analysis of human skin, offering unparalleled molecular specificity. Despite its promise, the pervasive issue of etalon artefacts—caused by multiple reflections and interference in back-illuminated charge-coupled devices (BI-CCDs)—significantly hinders its application, particularly in pigmented skin, which is prevalent in Asia. Addressing this critical challenge, we present a toward-automated digital processing workflow that combines independent component analysis (ICA) and continuous wavelet transform (CWT) to effectively identify and eliminate etalon artefacts. By leveraging the consistent fringe patterns generated in a dual-wavelength excitation Raman system, our method isolates the specific frequency and wavelength ranges impacted by etalon artefacts without requiring additional calibration or manual parameter tuning. Validation on in vivo Raman spectra from diverse skin samples demonstrates that this approach not only removes etalon artefacts with high precision but also preserves critical spectral features, establishing a robust foundation for accurate and detailed skin analysis.

Raman spectroscopy has become an essential tool in biomedical applications due to its ability to provide non-invasive, label-free, and highly sensitive detection of molecular compositions within biological tissues. Consequently, it has been increasingly utilized for skin analysis, enabling the detection of various chemical compounds that can offer insights into the physiological and pathological states of the skin, thereby contributing to skin condition monitoring and disease diagnosis [1–3]. Although Raman spectroscopy has proven to be a valuable tool for skin research, it faces several challenges that limit its application. The main issue is the etalon artefacts, which become more pronounced in weak signals, obscuring or distorting the faint spectral features. This challenge is particularly pronounced in pigmented skin, such as Asian skin with a higher melanin content. Importantly, this limitation has gained attention due to the expanding use of Raman spectroscopy in clinical dermatology and the growing emphasis on personalized skincare solutions tailored to diverse populations. Furthermore, as the Asian market emerges as one of the fastest-growing segments in the global skincare industry [4,5], addressing the etalon artefacts issue in Raman spectra for pigmented skin has become a pressing priority with considerable commercial and scientific implications.

Etalon artefacts arise from the Fabry-Pérot microcavity created by back-illuminated charge-coupled devices (BI-CCD) in the Raman spectrometer, visualizing as internal reflections and interference patterns [6–8]. Researchers have made significant efforts to eliminate the impact of these interference patterns and improve the accuracy and reliability of spectral data. Specifically, existing methods can be broadly categorized into two main approaches: physical design improvements and digital processing techniques. Physical design improvements have traditionally focused on optimizing the optical system to minimize or eliminate the etalon effect during signal acquisition. For example, Cvijin et al. suppressed etalon fringes by applying a low-amplitude audio-frequency vibration to a folding mirror in the laser cavity [9]. Kagawa et al. proposed a polarization control with a quarter-wave plate to reduce etalon fringes [10]. Although these methods suppress the generation of interference fringes, they also introduce significant drawbacks, including increased system complexity, higher costs, and more challenging operations, which impede commercialization. Moreover, physical modifications cannot completely eliminate etalon fringes. In contrast, digital processing techniques have gained popularity due to their flexibility and adaptability. The most straightforward strategy reported in previous work is to obtain a reference spectrum of the inherent background etalon pattern, either by measuring a standard sample or a calibration source before each acquisition, and subtracting it from the sample spectrum[11–14]. Conceptually, a broadband source could also serve this purpose. However, all reference-based approaches are limited by the fact that the etalon artefacts are time-variant, being influenced by temperature fluctuations and changes in system settings[15], and therefore require frequent recalibration. Another approach involves using mathematical models to fit the fringe pattern caused by the etalon effect as a baseline for correction[16–18], but these models often struggle to accurately describe the

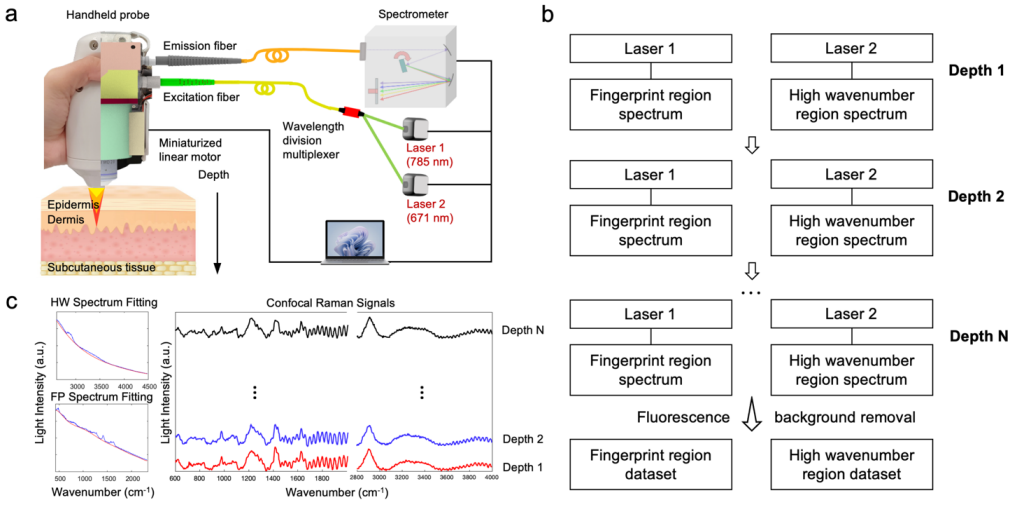


Figure 1. Experimental system and dataset preparation. (a) Experimental setup of simultaneous dual-wavelength source Raman spectroscopy with a handheld confocal probe. (b) Illustration of the dataset preparation method, where Raman signals are acquired through sequential dual-wavelength excitations. (c) Examples of measured Raman signals in the fingerprint (FP) and high-wavenumber (HW) regions, including directly measured signals and Raman signals at different depths after fluorescence background removal.

etalon artefacts and demand detailed knowledge of numerous system parameters. Filtering methods[19], deconvolution[20], wavelet threshold denoising[21,22], and nonlinear iterative partial least squares[23], have also been pass filtering is ineffective in isolating the etalon artefacts without degrading the signal. Additionally, these methods require careful selection of hyperparameters, such as filtering thresholds, employed to remove interference fringes in spectroscopy based on BI-CCD. However, the frequencies of the etalon fringes can overlap with those of the valid signal, making traditional band- wavelet decomposition levels, and the number of iterations, which are often determined empirically. This introduces variability and reliance on subjective judgment. With the advancement of deep learning, researchers have also explored neural networks for fitting etalon patterns or removing etalon fringes, achieving some success[24–26]. However, neural networks require large datasets for training to ensure robustness, demanding significant computational resources. Furthermore, supervised learning requires clean signals without etalon artefacts as labels, which are difficult to obtain. In summary, existing methods either require complex physical modifications or rely on computationally intensive digital techniques that often involve significant trade-offs in terms of efficiency, accuracy, and resource requirements. Therefore, there is a pressing need for a streamlined, robust, and automated method that does not require additional system measurements or extensive data collection.

We previously developed a dual-wavelength confocal Raman spectrometer (CRS) that simultaneously measures the same sample location with two different excitation wavelengths, thereby capturing both the fingerprint (FP) and high wavenumber (HW) regions [27,28]. This dual-wavelength excitation strategy is clinically motivated, offering a practical balance of cost, spectral coverage, and acquisition speed essential for non-invasive diagnostics. Importantly, the FP and HW spectra exhibit identical etalon fringe patterns because they traverse the same optical pathway and are acquired in strict synchrony. This inherent synchronization embeds a stable fixed-pattern reference directly within each measurement, thereby providing a self-contained basis for disentangling the artefacts from molecular signals. Independent component analysis (ICA) and continuous wavelet transform (CWT) offer complementary tools to achieve this separation. ICA [29,30] has gained widespread use as a powerful unsupervised learning method for solving blind source separation problems, as it reveals the underlying independent factors within a set of mixed signals. In Raman spectroscopy, the sample signals of various chemical components, etalon fringe patterns, and system random noise are statistically independent and non-Gaussian, which theoretically makes ICA a viable method for autonomously identifying etalon patterns. From a frequency analysis perspective, CWT originally developed for time-frequency analysis, has been extended as a two-dimensional analysis tool that can decompose frequency components along any measurement axis, including wavelength in spectral data [31,32]. In Raman spectroscopy, wavelet transform thus facilitates the identification of wavelength bands and frequency ranges impacted by etalon artefacts, offering a two-dimensional perspective that complements ICA-based source separation.

In this study, we developed a workflow combining the ICA and wavelet transform to identify and remove etalon artefacts based on the simultaneous dual-wavelength source Raman spectroscopy. Importantly, our approach does not impose any additional experimental or computational burden. Instead, it unlocks the spectral advantages already embedded in the dual-wavelength design, thereby advancing its analytical and clinical utility. To the best of our knowledge, this is the first toward-automated, model-free framework for etalon artefacts elimination that operates solely on the measured signals, without the need for additional calibration processes during acquisition. Validation on Raman spectra collected from various mixtures of amino acids and black ink and in vivo human skin demonstrated the method's effectiveness, laying a solid foundation for subsequent analysis. Compared to existing methods, our proposed method achieves significant SNR improvement while preserving spectral fidelity. With the increasing awareness of the challenges in optical sensing for Asian skin, the demand to eliminate etalon artefacts is going to grow rapidly. Our method holds significant relevance in this context and is expected to be widely adopted in Asian skin relevant applications.

MATERIALS AND METHODS

Experimental System

The experiment used the second-generation simultaneous dual-wavelength Raman spectroscopy system with a miniaturized handheld confocal Raman probe, which we previously developed [28]. The experimental setup is shown in Fig. 1(a). Its light source module is a fiber-coupled light input, consisting of two wavelengths (671 nm and 785 nm, 16.0 mW and 22.5 mW, respectively) laser beams, which would take only a single exposure to obtain both FP (450–1750 cm^{-1}) and HW (2800–3800 cm^{-1}) spectra. The laser power used in our system is significantly lower than those commonly used in similar in vivo confocal Raman systems[33,34], and is well within the safety limits specified by the ANSI Z136.1 standard for skin exposure. Due to the dual-wavelength confocal Raman system's capability to simultaneously excite the same sample depth with both 671 nm and 785 nm wavelengths, the Raman signals collected from both the FP and HW regions are acquired under identical environmental and system conditions. This design eliminates differential factors such as variations in optical alignment, temperature, or detector sensitivity that might otherwise affect the etalon fringes, ensuring consistent etalon patterns across both excitation wavelengths. Depositing the sample droplets onto the transparent film above the lens of the upward probe, or using the handheld probe to measure human skin. By advancing the lens to adjust the laser focal position, Raman signals were collected at N depths within the sample, as shown in Fig. 1(b). Each depth was measured with an exposure time of 5 µs, and the lens was advanced in step size of 10 µm. The examples of measured Raman signals and the fluorescence background removal datasets are shown in Fig. 1(c).

Signal Processing

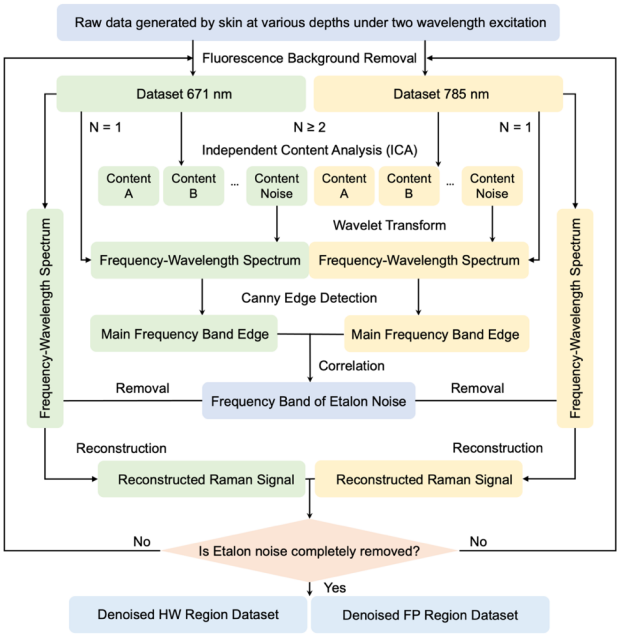


Figure 2. Workflow of ADWEEF for eliminating etalon artefacts based on a two-wavelength dataset.

To solve the effect of etalon artefacts in weak Raman signals, we propose a signal processing method, called Automated Dual-Wavelength etalon Elimination Framework (ADWEEF), based on the identical etalon patterns from dual-wavelength excitation as described in section A. The method combines ICA to identify the components containing etalon artefacts, with wavelet transform to obtain the frequency-wavelength spectrum. By analyzing the correlation between datasets from different excitation wavelengths, we identify the wavelength and frequency ranges affected by the common etalon pattern. Finally, inverse wavelet transform is used after thresholding to reconstruct the signal with the etalon artefacts removed. The following sections will provide a detailed explanation of the ICA, wavelet transform, and the overall workflow of the proposed method.

Independent Component Analysis (ICA). ICA is an unsupervised learning method used to separate mixed signals into underlying independent components. The core assumption of ICA is that the observed signals are linear mixtures of statistically independent, non-Gaussian source signals. Given a matrix of observed signals **X** (e.g., Raman signals), ICA models the observed signals as Eq. (1) [35]:

$$\mathbf{X} = \mathbf{AS} \tag{1}$$

where **A** is the unknown mixing matrix, and **S** is the matrix of independent source signals. ICA aims to find an unmixing matrix **W**, allowing **S** = **WX**. A critical condition for the successful application of ICA is that the number of observed signals must be greater than or equal to the number of independent components.

In this study, since the observed signals can be considered a linear combination of various components, ICA is applied to separate the statistically independent sources. The number of ICA components was selected as 5, corresponding to distinct spectral sources in both the FP and HW regions. In the FP region, the components include etalon artefacts, systematic noise, lipid, protein, and natural moisturizing factors. In the HW region, the components include etalon artefacts, systematic noise, water, lipid, and protein. These sources satisfy the assumptions of non-Gaussianity and statistical independence required by ICA, as verified in Supplementary Materials. Owing to the ability to acquire Raman spectra at multiple depths within the same sample, the proposed dual-wavelength confocal Raman system effectively provides the required independent measurements (d_1, d_2, …, d_n) for ICA, enabling the construction of an observation matrix.

Continuous Wavelet Transform (CWT). etalon artefacts is wavelength-dependent, meaning they only overlap with the valid Raman signals at specific wavelengths [7,17,36]. At other wavelengths, the etalon artefacts are too weak to distinguish from the valid signal. As a result, ICA may extract components that contain both etalon artefacts and parts of the Raman signals. CWT is a mathematical tool used to analyze non-stationary signals where the frequency content varies along the x-axis, providing a multi-scale frequency analysis. Its principle lies in decomposing a signal into wavelet functions at different scales and frequencies, enabling localized analysis. The CWT of a signal $x(w)$ is defined as Eq. (2):

$$W_x(a,b) = \int_{-\infty}^{\infty} x(w)\,\psi^*\left(\frac{w-b}{a}\right) dw \tag{2}$$

where $W_x(a,b)$ is the wavelet coefficient at scale a and translation b, $\psi(w)$ is the mother wavelet and ψ^* is the complex conjugate of the wavelet function. The Morlet wavelet was chosen as the mother wavelet due to its Gaussian-windowed complex sinusoidal form, which provides excellent time–frequency localization and is particularly well-suited for resolving quasi-

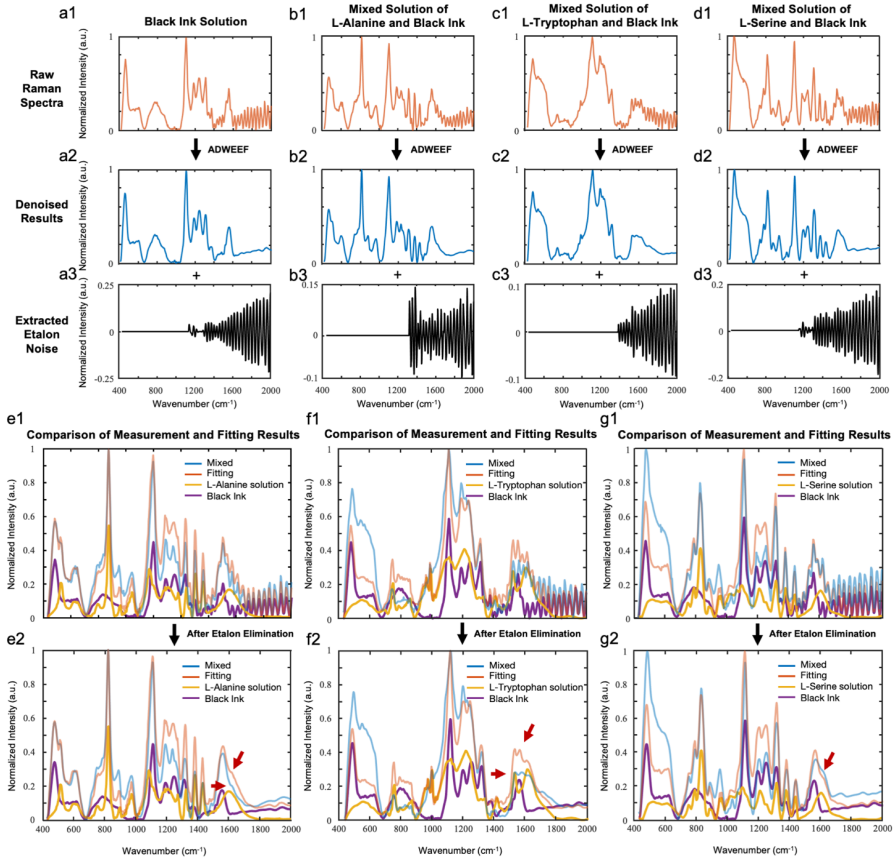


Figure 3. Results of etalon noise removal from Raman signals of chemical solutions mixed with black ink using our proposed method, ADWEEF, and comparison with the least squares decoupling for measured and denoised Raman signals. (a)(b)(c)(d) Panels a1, b1, c1 and d1 show the raw Raman spectra of the black ink solution and its mixed solution of L-Alanine, L-Tryptophan, and L-Serine, respectively. After etalon elimination by ADWEEF, panels a2, b2, c2, and d2 display the denoised results, and a3, b3, c3, and d3 show the extracted etalon noise from the raw data. (e)(f)(g) The top panels show the raw Raman signals of the three mixture solutions (blue lines) and their least squares fitting results (red lines) decoupled based on the raw Raman spectra of the amino acid solutions (yellow lines) and black ink (purple lines). The bottom panels show the fitting curves (red lines) derived from component ratios determined by decoupling previously, based on the raw Raman spectra of the amino acid solutions (yellow lines) and the etalon-removed Raman spectra of black ink (purple lines). The blue lines show the Raman spectra of the mixed solution after etalon elimination.

periodic features such as Raman peaks and broad periodic artifacts like etalon artefacts[37,38]. In this study, the x-axis represents wavelength. After applying CWT, the variation of frequency across different wavelengths will be obtained, enabling a clear visualization of the distribution of different frequency bands along the wavelength axis. The signal can be reconstructed from wavelet coefficients (modifying by thresholding to remove noise) by inverse continuous wavelet transform (iCWT) as Eq. (3):

$$x(w) = \int_0^\infty \int_{-\infty}^\infty W_x(a,b)\psi\left(\frac{w-b}{a}\right)\frac{dadb}{a^2} \qquad (3)$$

Workflow. The detailed workflow of ADWEEF (Fig. 2) for processing the Raman signals generated from samples at various depth under 671 nm and 785 nm excitation as follows. As illustrated in Fig. 1c, first, fluorescence backgrounds are removed via optimized polynomial fitting (sixth-order for FP region, fourth-order for HW region) following established protocols [39,40]. The preprocessed signals from N depths are compiled into wavelength-specific datasets for ICA. The specific component that includes the etalon pattern, called the etalon Independent Component (EIC), can be automatically identified by consistent weight coefficients sharing the same sign. Subsequent Morlet wavelet transforms convert these EICs into frequency-wavelength (F-W) spectra, where etalon patterns between the two datasets manifest as overlapping regions. And it can be extracted by the correlation between the resulting edges after applying a multi-scale Canny edge detection strategy (the initial σ value starts with a conservative 0.5, iteratively reduced by a factor of 10) to both spectra. Finally, by eliminating these regions from the original signal's F-W spectra and performing iCWT, the Raman spectra effectively removing the etalon artefacts will be successfully reconstructed. All processing was performed using MATLAB R2019b.

It is worth noting that our method is also applicable for removing etalon artefacts from Raman signals collected simultaneously at two excitation wavelengths within only one depth. In this situation, the F-W spectra of the Raman signals at the two excitation wavelengths are obtained separately using the wavelet transform. Subsequently, similar to the multi-depth case, the correlation analysis of the main frequency band edges is performed to identify the frequency band of etalon artefacts. Eliminating the signals within the identified frequency and wavelength ranges enables the reconstructed Raman spectra to be devoid of etalon artefacts.

Evaluation Matrix

The effective removal of etalon artefacts is demonstrated by eliminating interference fringes while preserving the original Raman spectral features. By analyzing the CWT results of a large number of Raman spectral signals (presented in section RESULTS AND DISCUSSION), we determined the wavelength range impacted by etalon artefacts is primarily concentrated after 868 nm (corresponding to 1220 cm^{-1} in the FP region and 3385 cm^{-1} in the HW region). Therefore, we use 868 nm as the threshold to partition the spectra into two regions: i. Region A ($\lambda < 868$ nm) is essentially unaffected by etalon artefacts. Here, the main concern is the fidelity of the reconstructed Raman features relative to the original spectra, which we quantified using the Root Mean Square Error (RMSE) as Eq. (4):

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(g_i - s_i)^2} \tag{4}$$

where, g_i represents the reconstructed signals, s_i the original signals, and N the number of spectral points.

ii. Region B ($\lambda \geq 868$ nm) is affected by etalon artefacts. In this region, the evaluation focuses on the effectiveness of artefact suppression, quantified using a region-based Signal-to-Noise Ratio (SNR). Specifically, the 868–905 nm range was treated as the *valid-signal band* because it still contains strong Raman peaks, while the >905 nm range was treated as the *noise band* dominated by etalon artefacts with negligible Raman contribution. The SNR was computed as Eq. (5):

$$\mathrm{SNR} = 20 \times \lg\left(\frac{A_{signal}}{A_{\mathrm{noise}}}\right) \tag{5}$$

where, A_{signal} is the root-mean-square amplitude of the normalized spectrum within the *valid-signal band*, and A_{noise} is the root-mean-square amplitude of the polynomial-fitted baseline within the *noise band*.

RESULTS AND DISCUSSION

Chemical Solution Results

To validate our method, we measured the Raman spectra of three representative amino acid aqueous solutions found in skin, L-Alanine, L-Tryptophan, and L-Serine (the yellow lines in Fig. 3(e1), (f1), and (g1)), using our simultaneous dual-wavelength source Raman spectroscopy with the confocal probe. Additionally, we employed black ink as a pigment-rich absorber phantom to mimic the weak-signal condition of melanin-rich skin. While it does not replicate melanin's scattering or fluorescence, its strong broadband absorption suppresses Raman signals and accentuates etalon fringes, thereby reproducing the key measurement challenge in vivo. We measured the Raman spectra of the three amino acid solutions with black ink (Fig. 3(b1), (c1), and (d1)), as well as for the black ink solution alone (Fig. 3(a1)). These Raman spectra were heavily influenced by pronounced etalon artefacts. We subsequently applied the ADWEEF proposed in Section MATERIALS AND METHODS (N = 1) for processing the Raman spectra of black ink and its mixtures with three different amino acids, as shown in Fig. 3(a2), (b2), (c2), and (d2). And the extracted etalon artefacts are shown in Fig. 3(a3), (b3), (c3), and (d3). Table 1 shows that the quantitative evaluation results demonstrate that the original Raman spectral characteristics of all solutions are well-preserved. At the same time, removing etalon artefacts significantly enhances the SNR. Given that the characteristic peaks of the amino acid solutions predominantly occur within the FP region, only the Raman signals excited at a wavelength of 785 nm are presented.

To analyze the spectral composition of the mixed solutions, we employed a linear least-squares fitting approach to determine the spectral contribution weights of each component (i.e., the amino acids solution and black ink) based on their individually measured spectra. This was done using the raw Raman spectra of the mixtures (blue lines in Fig. 3(e1), (f1), and (g1)). The component Raman spectra used in the fitting as references, amino acid solutions (yellow lines), and black ink (purple lines), were directly measured, without simulation or approximation. The resulting fitted spectra (red lines in Fig. 3(e1), (f1), and (g1)) reflect the theoretical spectral characteristics in the mixture. After removing etalon artefacts from the raw mixture spectra using ADWEEF, the denoised spectra are shown as blue lines in Fig. 3(e2), (f2), and (g2). The fitted spectra in Fig. 3(e2), (f2), and (g2) (red lines) were generated using the same weight coefficients obtained from the initial linear fitting step without re-fitting using etalon artefacts removed Raman spectra of black ink and the original amino acid solution spectra. The high degree of alignment between the denoised spectra and the fitted curves confirms the reliability of the ADWEEF process in recovering underlying spectral features. Notably, ADWEEF successfully restored small but critical Raman peaks previously buried in etalon artefacts, as highlighted by red arrows in Fig. 3. The recovery of these subtle spectral features is a critical advantage, as it enables the detection of hidden chemical signatures impacted by etalon artefacts, further validating the effectiveness of the proposed approach.

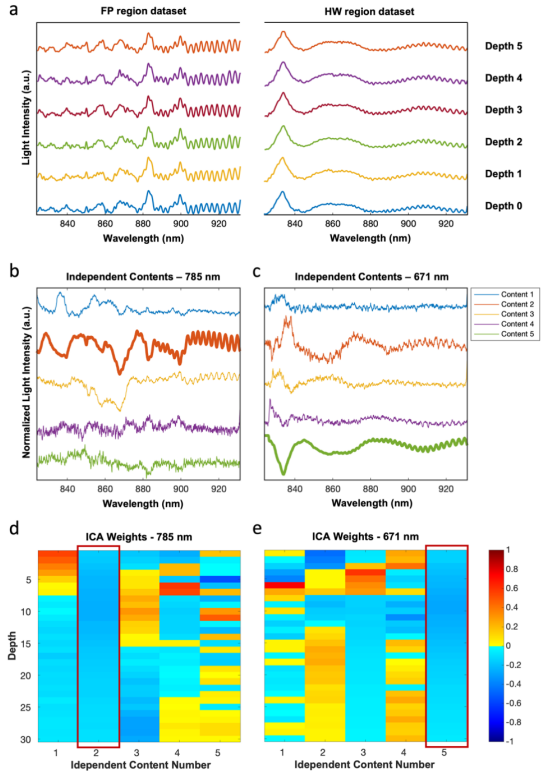


Figure 4. ICA Results of in vivo human skin. (a) Typical Raman signals in FP and HW region datasets with different depths. Results of independent component analysis (ICA) performed on datasets obtained from 785 nm (b) and 671 nm (c) excitations, and the corresponding weights of each independent component across signals from different depths (d-e).

In Vivo Human Skin Results

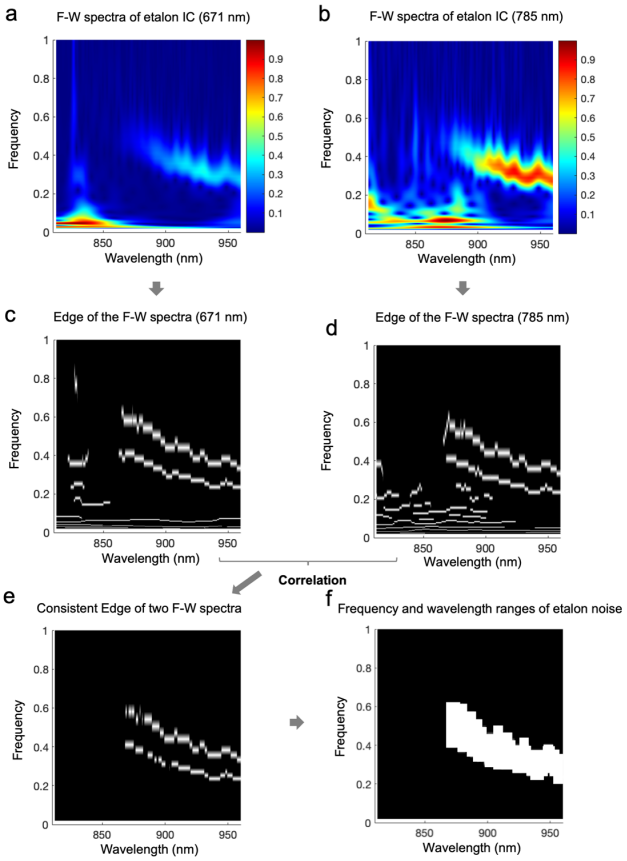


Figure 5. Frequency-Wavelength spectra of the etalon independent components identified from the datasets under 671 nm (a) and 785 nm (b) excitation, along with the edges extracted from them (c, d). (e) shows the correlation results of (c) and (d), representing the consistent edge in (c) and (d). (f) is filled inside the edge (e), representing the frequency and wavelength range affected by etalon artefacts.

For the further illustration of the feasibility of our proposed etalon artefacts elimination method, the in vivo human skin confocal Raman signal was collected with dual-wavelengths (671 nm and 785 nm) at 30 skin depths, and the fluorescence background was removed using polynomial fitting, as described before. Fig. 4(a) shows the typical Raman signals after background removal at six representative depths (including surface and deeper layers). Both the Raman signals in FP and HW regions still exhibited noticeable quasi-periodic noise, particularly in the latter part of the spectrum. This noise is what we call etalon artefacts, which affect the accuracy of subsequent analysis aimed at extracting biochemical component information from the Raman peaks. The five independent sources separated by ICA from the two datasets are shown in Fig. 4(b) and (c), respectively. And among them, the content including etalon fringes could be easily identified. For instance, as shown in Fig. 4(b) (content 2) and 4(c) (content 5), the components containing the etalon fringes exhibit the same sign (positive or negative) for weights in one dataset, as depicted in Fig. 4(d) and 4(e). While this analysis successfully reduces the influence of some noise factors, it is clear that the identified etalon-dominated component still contains residual Raman signals from biochemical compounds. This necessitates further analysis of this component using CWT to isolate the frequency regions affected by etalon fringes and enhance the extraction of the pure biochemical signal. Fig. 5(a) and 5(b) show the F-W spectra derived from EICs of datasets 671 nm and 785 nm by CWT. The consistent etalon artefact patterns, which manifest as high-frequency components in longer wavelength regions, were clearly identified in the spectrum. And the boundaries were detected by Canny algorithm subsequently, as shown in Fig. 5(c) and 5(d).

The correlation between the edges of the two wavelength datasets (671 nm and 785 nm) was used to extract the common etalon-dominated regions (Fig. 5(e)). After removing the wavelet coefficients in the wavelength and frequency ranges affected by etalon artefacts (Fig. 5(f)), the iCWT was applied to reconstruct the Raman signals for both the FP and HW regions at representative skin depths, with the results shown in Fig. 6(b), slightly smoothed for visual clarity. The extracted etalon artefacts shown in Fig. 6(c) are obtained by subtracting the denoised results (Fig. 6(b)) from the raw Raman spectra (Fig. 6(a)). As can be observed, the extracted etalon artefacts have no meaningful physiological signal or characteristic Raman peaks. Furthermore, contrary to assumptions in previous studies, the pattern of etalon artefacts cannot be accurately represented by a simple mathematical model (e.g., a sine model). Our proposed method, however, demonstrates superior performance in effectively extracting etalon artefacts and their elimination. In Fig. 6(d), the key Raman peaks revealed after etalon artefacts elimination are annotated, demonstrating the effectiveness of the proposed method in restoring physiologically relevant spectral features (the red dot line). Detailed vibrational mode assignments and molecular origins are provided in Supplementary Table S2.

Additionally, we compared the results denoised by our proposed method with other methods reported in previous etalon fringe elimination studies using two quantitative evaluation matrices [19–22]. A direct comparison and quantitative evaluation of our method against previously published approaches are presented in Fig. 7 and Table 2, respectively. The wavelength and frequency ranges affected by etalon artefacts, obtained through ICA followed by CWT analysis and correlation operations of EICs, were used as thresholds in iCWT and Fourier filtering, eliminating the need for multiple trials to determine appropriate thresholds. Both the visual comparison of curves (b) and (f), along with the SNR values, indicate that the performance of our method and Fourier filtering [19] in removing etalon artefacts is comparable. However, the Fourier method fails to isolate the wavelength regions affected by etalon artefacts during filtering, leading to a loss of signal detail in Region A, particularly for the FP region dataset (window i, ii). In contrast, our method preserves the Raman spectral features in both the HW and FP regions more effectively. The results of the wavelet threshold denoising (WTD) method [21], with automatic threshold detection, are shown in curve (c). While WTD achieves the lowest RMSE, it primarily excels at reducing random noise and is unable to recognize the quasi-periodic etalon artefacts, showing almost no change in SNR relative to the original signal (window iii). Since the etalon fringes are more likely to be superimposed on the effective Raman signal rather than convolved with it, deconvolution operations [20] cannot effectively remove them (window iv), as shown in curve (d). Complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) decomposes the signal into a series of intrinsic mode functions (IMFs), and WTD is then applied to remove noise components in the IMFs with low correlation [22]. The result is similar to WTD and is more suitable for removing high-frequency noise from the signal, as shown in curve (e). However, it still retains etalon fringes (win-

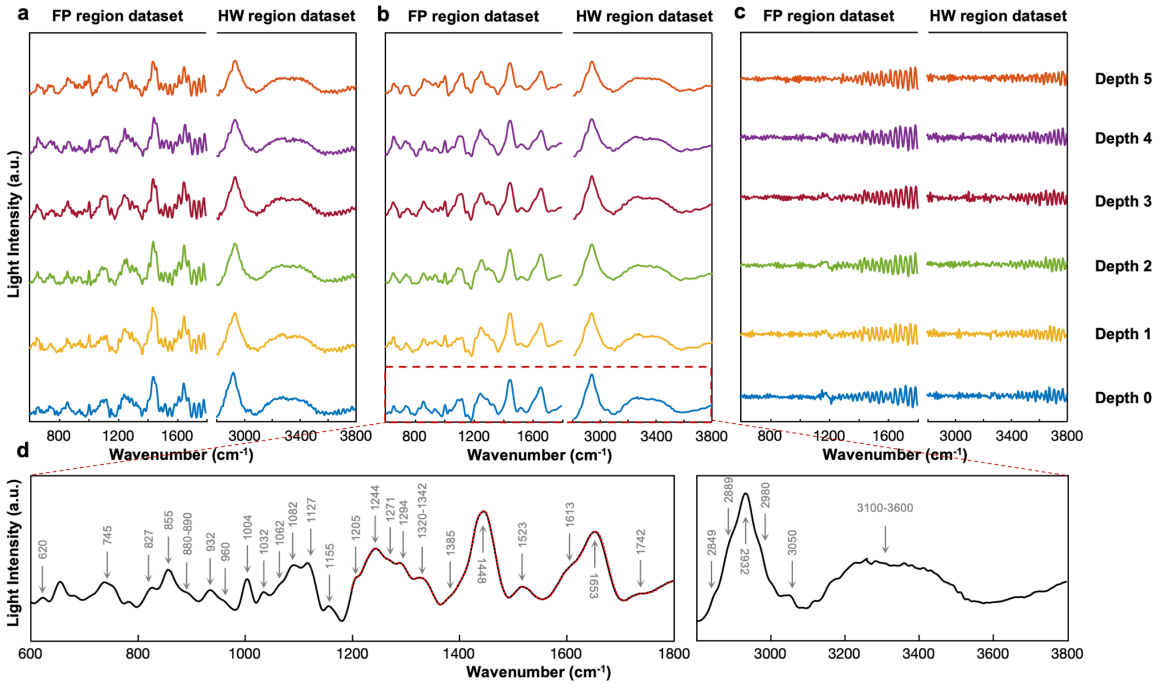


Figure 6. (a) The raw Raman spectra with etalon artefacts. (b) The etalon artefacts elimination results at typical depths reconstructed by the proposed ADWEEF. (c) The extracted etalon artefacts from the raw Raman spectra. (d) Raman peak assignment after etalon artefacts removal using the proposed ADWEEF. Detailed vibrational mode and molecular origin assignments are listed in Supplementary Table S2.

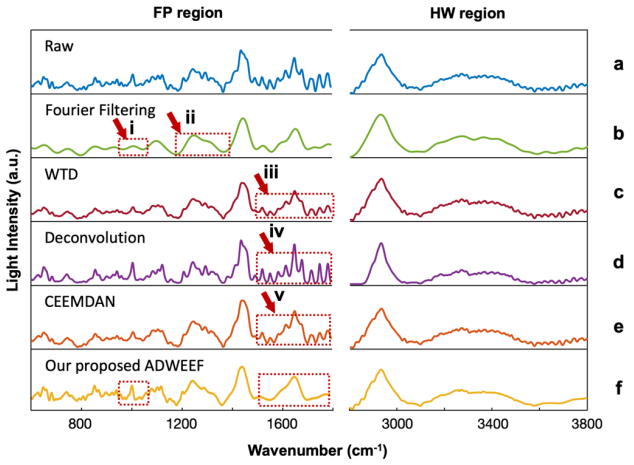


Figure 7. Comparison of raw Raman signals and processed signals after etalon artefacts removal using various methods. (a) represents the Raman signal with etalon artefacts after fluorescence background removal, while (b) - (f) show the signals after etalon artefacts elimination with Fourier filtering, wavelet threshold denoising (WTD), deconvolution, complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) combined with WTD, and our proposed method ADWEEF, respectively. Red annotations highlight key artefacts or failure modes in existing methods: (i)(ii) Significant peak distortion after Fourier filtering, where valid Raman peaks are overly smoothed; (iii) WTD primarily removes high-frequency noise but misinterprets etalon fringes as valid components, leaving them intact; (iv) In the deconvolution result, etalon artefacts is not suppressed and in some regions even amplified; (v) CEEMDAN also fails to fully remove fringe patterns, with residual periodic artefacts evident.

dow v), which interfere with the accurate identification and analysis of genuine Raman peaks originating from the skin.

By combining the visual results from Fig. 7 with the quantitative data in Table 2, it is clear that our method strikes a better balance between noise removal and signal preservation. Although other methods excel in certain quantitative metrics, our approach demonstrates greater robustness in accurately reconstructing the signal, particularly in preserving the finer details of the Raman spectra. In principle, a hybrid approach could be considered in the future. For example, WTD can be applied in Region A, and ADWEEF in more challenging areas, Region B. However, such integration would require further development to preserve full automation and signal continuity. Specifically, two key challenges remain: dynamically determining WTD thresholds without user input to keep the advantage of automation, and merging outputs from different methods without introducing discontinuities.

CONCLUSION

In conclusion, we developed a novel signal processing strategy based on the consistent etalon fringe patterns in the Raman spectra of FP and HW region collected with the previously established simultaneous dual-wavelength confocal Raman spectroscopy system. This method combines ICA and CWT to identify the specific wavelengths and frequency ranges affected by etalon artefacts, thereby effectively removing etalon artefacts from weak Raman signals. We validated the method on Raman spectra acquired from various mixtures of amino acids and black ink solutions, as well as from in vivo Asian facial skin samples. The results demonstrate that the etalon-induced distortions can be effectively eliminated, enabling the recovery of

Table 1. Evaluation Matrix of chemical solution results

Evaluation Matrix	Region A: Root Mean Square Error (RMSE)	Region B: Signal-to-Noise Ratio (SNR)	
FP Region		Original	Denoised
Black Ink	0.0237	6.2622	32.8459
L-Alanine	0.0120	8.8251	37.8708
L-Tryptophan	0.0127	8.0569	36.5966
Serine	0.0097	9.1329	34.8596

Table 2. Comparison of our proposed ADWEEF and other etalon elimination methods in previous study

Evaluation Matrix	Region A: Root Mean Square Error (RMSE)		Region B: Signal-to-Noise Ratio (SNR)	
Raman Region	FP Region	HW Region	FP Region	HW Region
Original Signal	/	/	9.9316	13.4676
ADWEEF	0.0223	0.0309	22.5938	24.8162
Fourier filtering	0.0641	0.0447	19.1288	29.2191
WTD	0.0237	0.0165	11.5165	14.7294
Deconvolution	0.0973	0.1511	8.8555	10.1046
CEEMDAN+WTD	0.0671	0.0245	13.5414	15.2271

previously obscured Raman peaks. It also demonstrated significant improvements in the SNR while preserving the original valid signals.

A key strength of this method lies in the use of a simultaneous dual-wavelength confocal Raman system. By leveraging the shared etalon patterns present in both the 671 nm and 785 nm wavelength signals, we eliminate the need for separate baseline measurements typically required in traditional methods. Moreover, the entire process is almost automated, reducing subjectivity and human error. The use of ICA, an unsupervised learning method, automatically identifies the independent components containing etalon fringes. By applying CWT to these components and correlating the results between the two wavelength datasets, filtering thresholds are determined dynamically, without the need for manual tuning. In our method, only two parameters remain: the number of ICA components, which can be chosen based on prior knowledge of the sample, and the initial Canny threshold, which requires adjustment only once for each system and sample type, thereby greatly reducing complexity compared with traditional approaches. This automated approach removes the potential bias and inconsistencies introduced by subjective parameter choices. Finally, a significant advantage of this method is its ability to avoid the pitfalls of traditional filtering techniques, which often rely on rigid rectangular thresholds that can unintentionally filter out valuable signal information. Through the use of CWT analysis, followed by edge detection and correlation, we can accurately identify and isolate the frequency-wavelength regions affected by the etalon pattern. This targeted filtering ensures that only the relevant noise components are removed, preserving the integrity of the Raman signals.

While the present method was developed and validated on a dual-wavelength Raman system, its underlying wavelet-domain principle can be extended to single-excitation configurations. The dual-excitation setup provides two spectra acquired in strict synchrony that share identical etalon fringe patterns while exhibiting distinct Raman features within the same wavelength domain due to the different excitation wavelengths. This spectral separation enables automatic identification of fringe-affected regions, facilitating a toward-automatic denoising process. In contrast, single-excitation systems lack this intrinsic reference, making fully automatic identification of etalon fringes more challenging. However, as demonstrated in our Supplementary Materials, we have also developed a CWT-based interactive GUI that allows manual selection of fringe-affected regions, providing a practical alternative for single-excitation data when partial user interaction is acceptable. Importantly, the dual-wavelength configuration remains clinically motivated, offering a practical balance of cost, spectral coverage, and acquisition speed that aligns well with the needs of real-time or intraoperative diagnostics. With the added benefit of providing richer chemical information, dual-wavelength Raman spectroscopy has been increasingly adopted in biomedical applications [41,42]. Our method builds upon this configuration without introducing additional optical complexity, instead providing an automated processing framework that exploits the shared etalon structure for effective denoising, eliminates the need for separate baseline measurements, and ensures robust fringe suppression without external reference data.

This method shows particular potential in improving the identification and extraction of weak Raman signals that are often masked by etalon artefacts. By accurately distinguishing between noise and signal, this approach enables more reliable detection of subtle biochemical changes, which is essential for applications involving low-intensity signals, such as pigmented skin spectroscopy. Moreover, the proposed method can be extended to any BI-CCD-based dual- and multi-wavelength systems, offering great potential in a wide range of applications such as food science[43] and chemical measurement[44]. This is because etalon artefacts originate from the multiple internal reflections of near-infrared photons within the photosensitive silicon layer of the BI-CCD sensor, due to the large refractive index mismatch between the silicon and the insulating layer. As a result, the formation of etalon fringes is independent of sample

properties and thus broadly generalizable. A more detailed explanation of this mechanism is provided in the Supplementary Materials.

SUPPORING INFORMATION

Mechanism of Etalon Artefacts Generation in Back-Illuminated CCD Systems (Section I); Validation of ICA Assumptions on Raman Spectral Components (Section II); Representative Raw Raman Spectra and Fluorescence Background Removal Process (Section III); Raman Peak Assignment of Human Skin after Etalon Artefacts Elimination (Section IV); Trend of Research on Dual- and Multi-wavelength Raman System (Section V); CWT-based Interactive GUI for Single-Excitation Etalon Removal (Section VI). The codes related to this work have been made available on GitHub at: https://github.com/Dolores99/ADWEEF-etalon-removal

AUTHOR INFORMATION

Corresponding Author

* Malini Olivo, Email: malini_olivo@asrl.a-star.edu.sg,
* Steven Tien Guan Thng, Email: steventhng@nsc.com.sg,
* Qian Cheng, Email: q.cheng@tongji.edu.cn.

Author Contributions

All authors have given approval to the final version of the manuscript.

ACKNOWLEDGMENT

This study was supported by A*STAR I&E Gap fund [grand number I24D1AG001], A*STAR BMRC CRF fund 2024, and Human Health and Potential (HHP) IAF-pp project OPTIMAS H24J4a0146.

REFERENCES

(1) Kanemura, Y.; Kanazawa, M.; Hashimoto, S.; Hayashi, Y.; Fujiwara, E.; Suzuki, A.; Ishii, T.; Goto, M.; Nozaki, H.; Inoue, T.; Takanari, H. Assessment of Skin Inflammation Using Near-Infrared Raman Spectroscopy Combined with Artificial Intelligence Analysis in an Animal Model. Analyst 2022, 147 (12), 2843–2850.

(2) Franzen, L.; Windbergs, M. Applications of Raman Spectroscopy in Skin Research — From Skin Physiology and Diagnosis up to Risk Assessment and Dermal Drug Delivery. Adv Drug Deliv Rev 2015, 89, 91–104.

(3) Feng, X.; Moy, A. J.; Nguyen, H. T. M.; Zhang, J.; Fox, M. C.; Sebastian, K. R.; Reichenberg, J. S.; Markey, M. K.; Tunnell, J. W. Raman Active Components of Skin Cancer. Biomed Opt Express 2017, 8 (6), 2835.

(4) Baek, E.; Lee, H. J.; Oh, G. E. Understanding East Asian Consumers' Responses to Inclusive Beauty Products in Advertising. Int J Advert 2023, 42 (5), 868–889.

(5) Krishnan, S.; Safia Amira, N.; Nur Atilla, U.; Syafawani, S.; Hafiz, M. The Usage of Cosmetic in Malaysia: Understanding the Major Factors That Affect the Users. Management 2017, 2017 (1), 48–51.

(6) Werle, P. Accuracy and Precision of Laser Spectrometers for Trace Gas Sensing in the Presence of Optical Fringes and Atmospheric Turbulence. In Applied Physics B: Lasers and Optics; 2011; Vol. 102, pp 313–329.

(7) Groom, D. E.; Holland, S. E.; Levi, M. E.; Palaio, N. P.; Perlmutter, S.; Stover, R. J.; Wei, M. Quantum Efficiency of a Back-Illuminated CCD Imager: An Optical Approach. In Sensors, Cameras, and Systems for Scientific/Industrial Applications; Blouke, M. M., Williams, Jr., G. M., Eds.; SPIE: California, 1999; pp 80–90.

(8) Mitra, C. Mid-Infrared Spectroscopy and Challenges in Industrial Environment. In Infrared Spectroscopy - Principles, Advances, and Applications; Marwa El-Azazy, Ed.; IntechOpen, 2019.

(9) Vujkovic Cvijin, P.; Wells, W. K.; Gilmore, D. A.; Wu, J.; Hunten, D. M.; Atkinson, G. H. Fringe Pattern Suppression in Intracavity Laser Spectroscopy. Appl Opt 1992, 31 (27).

(10) Kagawa, N.; Wada, O.; Koga, R. Suppression of the e ′ Talon Fringe in Absorption Spectrometry with an Infrared Tunable Diode Laser. Optical Engineering 1997, 36 (9), 2586–2592.

(11) Nikodem, M.; Gomółka, G.; Klimczak, M.; Pysz, D.; Buczyński, R. Demonstration of Mid-Infrared Gas Sensing Using an Anti-Resonant Hollow Core Fiber and a Quantum Cascade Laser. Opt Express 2019, 27 (25), 36350.

(12) Slemr, F.; Harris, G. W.; Hastie, D. R.; Mackay, G. I.; Schiff, H. I. Measurement of Gas Phase Hydrogen Peroxide in Air by Tunable Diode Laser Absorption Spectroscopy. Journal of Geophysical Research: Atmospheres 1986, 91 (D5), 5371–5378.

(13) Massie, C.; Chen, K.; Berger, A. J. Calibration Technique for Suppressing Residual Etalon Artifacts in Slit-Averaged Raman Spectroscopy. Appl Spectrosc 2022, 76 (2), 255–261.

(14) Werle, P.; R., M.; F., S. The Limits of Signal Averaging in Atmospheric Trace-Gas Monitoring by Tunable Diode-Laser Absorption Spectroscopy (TDLAS). Applied Physics B 1993, 57, 131–139.

(15) Cao, J.; Wang, Z.; Zhang, K.; Yang, R.; Wang, Y. Etalon Effects Analysis in Tunable Diode Laser Absorption Spectroscopy Gas Concentration Detection System Based on Wavelength Modulation Spectroscopy. In 2010 Symposium on Photonics and Optoelectronics; IEEE, 2010; pp 1–5.

(16) Zhang, E. J.; Tombez, L.; Teng, C. C.; Wysocki, G.; Green, W. M. J. Adaptive Etalon Suppression Technique for Long-term Stability Improvement in High Index Contrast Waveguide-based Laser Absorption Spectrometers. Electron Lett 2019, 55 (15), 851–853.

(17) Hu, B. L.; Zhang, J.; Cao, K. Q.; Hao, S. J.; Sun, D. X.; Liu, Y. N. Research on the Etalon Effect in Dispersive Hyperspectral VNIR Imagers Using Back-Illuminated CCDs. IEEE Transactions on Geoscience and Remote Sensing 2018, 56 (9), 5481–5494.

(18) Schnell, U.; Gray, S.; Dändliker, R. Dispersive White-Light Interferometry for Absolute Distance Measurement with Dielectric Multilayer Systems on the Target. Opt Lett 1996, 21 (7), 528.

(19) Sun, H. C.; Whittaker, E. A. Novel Etalon Fringe Rejection Technique for Laser Absorption Spectroscopy. Appl Opt 1992, 31 (24), 4998–5002.

(20) Pistorius, A. M. A.; DeGrip, W. J. Deconvolution as a Tool to Remove Fringes from an FT-IR Spectrum. Vib Spectrosc 2004, 36 (1), 89–95.

(21) Li, C.; Guo, X.; Ji, W.; Wei, J.; Qiu, X.; Ma, W. Etalon Fringe Removal of Tunable Diode Laser Multi-Pass Spectroscopy by Wavelet Transforms. Opt Quantum Electron 2018, 50 (7).

(22) Wang, S.; Gong, W.; Wang, Z.; Wei, Y.; Li, Y.; Zhang, T.; Zhang, Q.; Zhang, L.; Song, F.; Zhang, W.; Liu, T. Interference Fringe Suppression in Tunable Diode Laser Absorption Spectroscopy Based on CEEMDAN-WTD. Front Phys 2022, 10.

(23) Hren, B.; Mink, J.; Balázs, L. Fourier Transform Infrared Analysis of Gas Composition in Low-Volume Light Bulbs. Anal Chem 2002, 74 (24), 6402–6407.

(24) Nicely, J. M.; Hanisco, T. F.; Riris, H. Applicability of Neural Networks to Etalon Fringe Filtering in Laser Spectrometers. J Quant Spectrosc Radiat Transf 2018, 211, 115–122.

(25) Huang, C.; Zhang, T.; Kong, X.; Li, Y.; Wei, H. Deep-Learning-Enabled High-Fidelity Absorbance Spectra from Distorted Dual-Comb Absorption Spectroscopy for Gas Quantification Analysis. Appl Spectrosc 2024, 78 (3), 310–320.

(26) Luo, Q.; Zhou, J.; Li, W.; Yang, C.; Gui, W. Interference Fringe Suppression for Oxygen Concentration Measurement Using Adaptive Harmonic Feeding Generative Adversarial Network. IEEE Sens J 2022, 22 (3), 2419–2429.

(27) Zhang, R.; Bi, R.; Ho Jun Hui, C.; Rajarahm, P.; Dinish, U. S.; Olivo, M. A Portable Ultrawideband Confocal Raman Spectroscopy System with a Handheld Probe for Skin Studies. ACS Sens 2021, 6 (8), 2960–2966.

(28) Qi, Y.; Zhang, R.; Rajarahm, P.; Zhang, S.; Ebrahim Attia, A. B.; Bi, R.; Olivo, M. Simultaneous Dual-Wavelength Source Raman Spectroscopy with a Handheld Confocal Probe for Analysis of the

Chemical Composition of In Vivo Human Skin. Anal Chem 2023, 95 (12), 5240–5247.

(29) Comon, P. Independent Component Analysis, A New Concept? Signal Processing 1994, 36 (3), 287–314.

(30) Chung, S. H.; Park, C. S.; Park, K. S. Application of Independent Component Analysis (ICA) Method to the Raman Spectra Processing. In Optical Diagnostics and Sensing V; SPIE, 2005; Vol. 5702, p 168.

(31) Daubechies, I. The Wavelet Transform, Time-Frequency Localization and Signal Analysis. IEEE Trans Inf Theory 1990, 36 (5), 961–1005.

(32) Hoang, V. D. Wavelet-Based Spectral Analysis. TrAC Trends in Analytical Chemistry 2014, 62, 144–153.

(33) Wu, Z.; Jiang, L.; Wang, W.; Zhao, J.; Lui, H.; Zeng, H. Precise in Vivo Tissue Micro-Raman Spectroscopy with Simultaneous Reflectance Confocal Microscopy Monitoring Using a Single Laser. Opt Lett 2019, 44 (6), 1383.

(34) Duvic, M.; Ni, X.; Talpur, R.; Herne, K.; Schulz, C.; Sui, D.; Ward, S.; Joseph, A.; Hazarika, P. Tazarotene-Induced Gene 3 Is Suppressed in Basal Cell Carcinomas and Reversed In Vivo by Tazarotene Application. Journal of Investigative Dermatology 2003, 121 (4), 902–909.

(35) Shlens, J. A Tutorial on Independent Component Analysis. 2014.

(36) Hult, J.; Watt, R. S.; Kaminski, C. F. Dispersion Measurement in Optical Fibers Using Supercontinuum Pulses. Journal of Lightwave Technology 2007, 25 (3), 820–824.

(37) Le, T.-P. Use of the Morlet Mother Wavelet in the Frequency-Scale Domain Decomposition Technique for the Modal Identification of Ambient Vibration Responses. Mech Syst Signal Process 2017, 95, 488–505.

(38) Shao, L.; Lin, X.; Shao, X. A Wavelet Transform and Its Application to Spectroscopic Analysis. Applied Spectroscopy Reviews. 2002, pp 429–450.

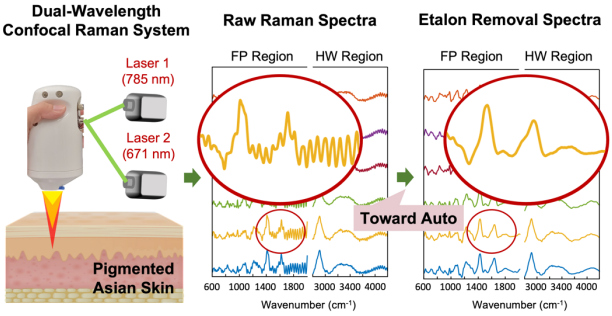
(39) Kiselev, R.; Schie, I. W.; Aškrabić, S.; Krafft, C.; Popp, J. Design and First Applications of a Flexible Raman Micro-Spectroscopic System for Biological Imaging. Biomed Spectrosc Imaging 2016, 5 (2), 115–127.

(40) Cadusch, P. J.; Hlaing, M. M.; Wade, S. A.; McArthur, S. L.; Stoddart, P. R. Improved Methods for Fluorescence Background Subtraction from Raman Spectra. Journal of Raman Spectroscopy 2013, 44 (11), 1587–1595.

(41) Kiefer, J. Dual-Wavelength Raman Fusion Spectroscopy. Anal Chem 2019, 91 (3), 1764–1767.

(42) Masson, L. E.; O'Brien, C. M.; Pence, I. J.; Herington, J. L.; Reese, J.; van Leeuwen, T. G.; Mahadevan-Jansen, A. Dual Excitation Wavelength System for Combined Fingerprint and High Wavenumber Raman Spectroscopy. Analyst 2018, 143 (24), 6049–6060.

(43) Ye, J.; Li, J.; Lu, M.; Qi, X.; Li, B.; Wei, H.; Li, Y.; Zou, M. Dual-Wavelength Excitation Combined Raman Spectroscopy for Detection of Highly Fluorescent Samples. Appl Opt 2021, 60 (23), 6918.

(44) Lee, J. H.; Jeong, Y. S.; Koh, Y. jin; Kim, J.; Nam, H.; Son, H.; Choi, S. K. Detection of Hazardous Chemical Using Dual-Wavelength Raman Spectroscopy in the Ultraviolet Region. Spectrochim Acta A Mol Biomol Spectrosc 2023, 287, 122061.

For Table of Contents Only