Automatic morphology-based cubic p-spline fitting methodology for smoothing and baseline-removal of Raman spectra

Juan José González-Vidal[a,b*], Rosanna Pérez-Pueyo[a], María José Soneira[a]
[a] Signal Theory and Communications Department, ETSETB. Universitat Politècnica de Catalunya. C/ Sor Eulàlia d'Anzizu s/n, D5, Campus Nord, 08034, Barcelona, Spain. E-mail: {juanjo.glezvidal, rperez, soneira}@tsc.upc.edu
[b] DAPCOM Data Services, European Space Agency Business Incubation Centre Barcelona, Parc UPC-RDIT, c/ Esteve Terrades 1, 08860 Castelldefels, Barcelona, Spain
[*] **Corresponding author**: Juanjo González-Vidal. Address: Signal Theory and Communications Department, ETSETB. Universitat Politècnica de Catalunya. C/ Sor Eulàlia d'Anzizu s/n, D5, Campus Nord, 08034, Barcelona, Spain. Telephone: +34 934016442. Fax: +34 934016447. E-mail: juanjo.glezvidal@tsc.upc.edu

Abstract

Noise filtering is considered a crucial step for the proper interpretation of Raman spectra. In this work, we present a new denoising procedure which enhances the Raman information while reducing unwanted contributions from the most frequent noise sources, i.e. the shot noise and the fluorescence's baseline. The procedure increases the signal-to-noise ratio while preserving simultaneously the shapes, positions and intensity ratios of the Raman bands. The method relies on cubic p-spline fitting and mathematical morphology, and requires no user input. We describe the details of this method and include a benchmark to study the performance of the presented approach compared to the most commonly used denoising techniques. The method has been successfully applied to improve the signal quality of Raman spectra from artistic pigments. The reliable results that were obtained make the methodology a useful tool to help the analyst in the interpretation of Raman spectra from pigments in art works.

Keywords: Raman spectroscopy; noise filtering; shot noise; fluorescence's baseline; pigment analysis

Introduction

Raman spectroscopy is extensively used to analyze the composition and structure of a wide range of organic and inorganic materials in a non-destructive fashion. A Raman spectrum may provide valuable information about the analysed sample. Nonetheless, the quality of this information may be compromised due to the presence of interfering or unwanted signals. Broadly speaking a Raman spectrum can be divided into two parts: the useful signal and the noise[1]. In our case, the useful signal is the Raman information, which can be seen as a fingerprint signal in the form of a specific combination of peaks –the Raman bands– by which the analyzed material can be unequivocally identified. In contrast, the noise is the part of the Raman spectrum that comes from undesired sources, which thus can adversely affect the interpretation of the analyzed sample.

The most commonly found sources of noise in Raman spectroscopy are shot noise and fluorescence's baseline: the shot noise is an unavoidable noise source caused by the statistical nature of light, which may compromise the analysis of a Raman spectrum; the fluorescence's baseline is sample-inherent usually of higher amplitude than the Raman information that can thus mask the Raman bands. Therefore, the noise impact should be reduced –i.e. filtered– before performing further analyses (whether through visual inspection or automated methodologies) in order to accomplish a proper interpretation of a Raman spectrum.

There is no single strategy for noise filtering in Raman spectroscopy. Several methods have been proposed to enhance the Raman information[2-24].The most frequently used methods are software procedures, which do not require to upgrade the existing instrumentation. Such procedures are generally dedicated to filter one kind of noises separately, i.e. or shot noise or fluorescence's baseline. For instance, to reduce the shot noise the simplest procedure is the median filter, whilst to remove the fluorescence's baseline the simplest and widest used method is the polynomial fitting. The basic version of such methods involves user intervention in order to select appropriate key parameters, and this selection process is usually time consuming. For instance, choosing which Raman shifts belong to noise sources in non-Raman characteristic band regions or which ones belong to Raman characteristic band regions is a critical point, which may introduce subjectivity depending on the analysts' experience. Thus, several methods have been developed in the last decade in order to avoid any user intervention. Generally, such methods are based on iterative solutions. Though these methods may provide successful results, they treat one kind of noise at a time and due to the high nonlinearity and complexity of a Raman spectrum they may not well smooth it or reject its fluorescence's baseline. As an alternative, the fully-automated noise filtering approach developed in this research pursues a twofold objective: the shot noise reduction and the fluorescence's baseline removal.

In this respect, this paper introduces a new and simple procedure to reduce the shot noise and to remove the fluorescence's baseline simultaneously with a single strategy, which is independent of the Raman spectrum to be filtered. The underlying principle of this novel approach is based on the different "shapes" shown by the shot noise and the fluorescence's baseline in a Raman spectrum: the shot noise may be seen as an intensity fluctuation (rapid variation), whilst the shape of the fluorescence background is shown as a soft drifting baseline (slow variation). In this regard, the method uses mathematical morphology operations, which simplify and preserve the main features of the shapes, jointly with cubic penalized spline fitting for smoothing and baseline-removal of Raman spectra in a unified way. No parameter tweaking is needed and therefore no user intervention is required. The method was developed as an application-specific algorithm which improves the signal-to-noise ratio tackling at the same time both shot noise and baseline rejection, preserving the shapes, positions and intensity ratios of the Raman bands.

In particular, the denoising methodology developed in the current work has been successfully applied to the analysis of works of art. There is a large international consensus that cultural heritage must be conserved and preserved for future generations[25-29]. Hence, a thorough knowledge of the pigments present in an art work is absolutely essential to gain insight into the materials composition and deterioration mechanisms in order to apply optimum restoration and conservation methodologies[30-33]. In this sense, the Raman spectra of artistic pigments are specific to the vibrational modes of the molecules of an analyzed work of art, property that gives to Raman spectroscopy a large potential for identifying pigments. Nevertheless, pollutants and other environmental factors, as well as interferences from the binding media and to aging, may have a direct impact on the quality of the Raman signal[34-37], which contributes to the difficulties in identifying pigments by Raman spectroscopy in the form of noise. An appropriate signal treatment expands the capabilities of the technique to non-invasively identify and quantify the chemical composition of paint layers in art works.

The methodology presented here describes the core principles of the proposed approach for noise filtering. Then, we present a benchmark for the evaluation and comparison of the performance of the proposed filter and the most widely used noise filtering techniques using simulated spectra. Finally, the results are discussed and evaluated for real-case examples.

Methodology

Noise is an intrinsic factor in Raman spectroscopy and negatively affects the Raman information. Accordingly, noise filtering is a preliminary process decisive in the analysis of Raman spectra. In this regard, a filter methodology is proposed which broadly speaking is based on a curve fitting technique intended to obtain an improved signal-to-noise ratio (SNR), making the Raman spectra easier to interpret.

The use of piecewise polynomials to model regression functions and perform curve fitting has a long history[38-44]. In smoothing, the location of the points, or *knots*, in which the polynomial pieces are joined are arbitrary which permits a very large class of possible fits. A widely used fit is based on splines[45-48], which are piecewise-defined by polynomial functions. Penalized splines (or p-splines)[49] are a very popular spline fitting approach, which has the following properties: efficient computation, flexibility, and ease of setup[50]. P-splines are regression splines fit by least-squares with a roughness penalty which avoids overfitting[46]. According to[51], the optimal degree of this piecewise polynomial regression is 3, which generates the so-called *cubic p-splines*. The smoothness of the estimate varies as a function of the smoothing parameter, λ: the larger the smoothing parameter, the more the fit minimizes towards a polynomial fit, which in turn allows the estimate to deal with data gaps[49]. In our research, the λ-parameter was selected to be small enough so as to keep the estimates smooth and its value was fixed to 0.7. This constant value provides a good compromise between smoothness and polynomial fit regardless of the input spectra, whether simulated or experimental.

The choice of knots has been a subject of much research[52,53]. Equidistant knots can be used, but this allows only limited control over the fit. Instead, a smart knots selection is preferred so in the case noise filtering of Raman spectra the presence of noise is optimally reduced whilst the shape and positions of the Raman bands remain unaltered. This may be achieved through a strategic selection of knots according to the shape of the input data. To do so, the usage of mathematical morphology operations is proposed in the current work.

Mathematical morphology is a nonlinear technique based on classical set theory[54,55]. It finds application in many different research fields as it only involves the definition of sets of data taking advantage of the properties of those sets[56,57]. In particular, it is predominantly useful in fields in which the shape is the most important feature. Morphological operations transform the original function into another function looking for geometric structures (i.e. shapes) using the structuring element whose shape is chosen according to the "morphology" of the function and the special structures to be extracted. Choosing a suitable structuring element, we can use the information extracted from morphological operations to generate the knots sequences to be used in the cubic p-splines fitting to filter a noisy Raman spectrum. There are two basics operations in mathematical morphology, called erosion and dilation and the combination of such operations provides two more operators named closing and opening. The morphological closing of a function *f* by a structuring element *Y*, $\phi_Y(f)$, is described mathematically as $\phi_Y(f) = \varepsilon_Y[\delta_Y(f)]$, being $\varepsilon_Y(f)(x) = \min_{s \epsilon Y} f(x+s)$ the erosion and $\delta_Y(f)(x) = \max_{s \epsilon Y} f(x+s)$ the dilation of the function. The closing smoothes the function nonlinearly removing holes and connecting nearby items thus taking always values that are higher or equal than those of the input function. Hence, the closing by a short structuring element may provide a rough estimation the shape of the Raman bands. In this case, this short structuring element, *Ymin*, is defined so that the closing can take into account any Raman band, and therefore fixed to three data points. The resulting closing is modified to further reduce the shot noise influence as $\phi'_{Ymin}(f) = \phi_{Ymin}(f) \notin \varepsilon_{Ymin}(f)$. On the other hand, the morphological opening of a function *f* by a structuring element *Y*, $\gamma_Y(f) = \delta_Y[\varepsilon_Y(f)]$, smoothes the input function too but differently since it removes the positive peaks, taking always values that are lower or equal than those of the input function. Hence, the opening by the optimal structuring element, *Yopt*, may provide a rough estimation of the fluorescence's baseline. This optimal structuring element, is selected following a lookup procedure[14]:

i) As starting point, the opening of the input spectra by the minimum structuring element is computed
ii) Iteratively, the opening by an incremented length of the structuring element is computed for each iteration
iii) The root mean square error (RMSE) between the resulting opening and the opening of the previous iteration is computed
iv) The optimal structuring element is obtained when the RMSE gets stabilized, i.e. the same opening is obtained in 3 consecutive iterations
v) The resulting opening is modified in order to reduce any flaw in the peak regions as

$$\gamma'_{Yopt}(f) = min(\gamma_{Yopt}(f), \frac{\varepsilon_{Yopt}(f) + \delta_{Yopt}(f)}{2})$$

The methodological scheme of the noise filtering presented in the current work follows the flowchart shown in Fig. 1. Being *f* a noisy Raman spectrum, a knots sequence, K1, is obtained from the intersection of f with the modified closing by the minimum structuring element, $\phi'_{Ymin}(f)$. A cubic p-spline fit of *f* through K1 is performed which provides an intermediate function, g. Then, the optimal structuring element that follows the morphology of the baseline in g is achieved. Next, a new knots sequence, K2, is obtained from the intersection of g with the modified opening by the optimal structuring element, $\gamma'_{Yopt}(g)$. A cubic p-spline fit of g through K2 provides an estimation of the fluorescence's baseline, h. Finally, the denoised spectrum, *d*, is obtained by subtracting the fluorescence's baseline estimation from the intermediate function.

The morphology-based cubic p-spline fitting methodology for enhancing Raman spectra is graphically represented in Fig. 2, where it was applied to a measured Raman spectrum of a sample of a PY1 pigment powder.

Results and Discussion

Analysis on simulated Raman spectra

The proposed filtering methodology was tested using simulated spectra. In this sense, a simulated spectrum was generated by combining a variable number of Lorentzian bands with random locations, amplitudes and FWHM, constrained such that it appeared qualitatively similar to real Raman spectra. Shot noise was simulated from a zero-mean Gaussian distribution and variable variance. Also, different artificial profiles were simulated to mimic the fluorescence's baseline, which were selected in a heuristic way but similar in appearance to that of Raman spectra. In particular, four simulated profiles were used: polynomial, linear, sigmoidal and sinusoidal.

Comparisons were performed with the here proposed filtering approach and several techniques in common use to filter the shot noise, such as the Wiener filter[58], the median filter, the wavelet filter[59], the FFT filter, and the fuzzy filter previously developed by the authors[4]. For the Wiener filter the noise was estimated from the ideal spectra, and the response function was used with no smearing. The median filter was run several times with window sizes ranging from 3 to 11 data points and the window providing the lowest RMSE between the denoised and the ideal spectra was selected. The wavelet filter was performed by means of the standard wavelet soft thresholding with default parameters. The FFT filter was run several times with rectangular filters of different sizes, selecting the one providing the lowest RMSE. Additionally, comparisons with respect to baseline rejection were carried out with the here proposed filtering methodology, the morphology-based filtering approach published in[14] and the conventional polynomial approach, being the last one the most popular method in Raman spectroscopy for subtracting the fluorescence's baseline. The conventional polynomial method was run several times selecting the polynomial degree that provided the lowest RMSE.

Unlike the here proposed methodology, the general techniques in common use here tested are focused on either shot noise filtering or baseline rejection. Therefore, to perform a proper comparison with respect to the filtering approach presented in the current paper, the previously commented shot noise filtering techniques were combined with the baseline filters above-mentioned. In particular, 100 noisy spectra were simulated and the RMSE between the ideal and the filtered spectra was computed to compare the results of the here proposed approach and each of the combinations of a shot noise filtering technique with a baseline filter. The results, i.e. mean RMSE and standard deviation, are shown in Table 1 – the best-degree polynomial filter is represented as *PF* and the morphology-based filter[14] is represented as *MF*. On average, at the noise levels tested the here presented method outperforms the combination of the other techniques. From the results we may also say that the proposed filter provides the least distortion of the Raman bands, which is very useful when the spectrum must be subsequently processed in order to identify the material or quantify its proportion in mixtures. Additional test results using simulated spectra can be found in the supplementary material (Fig. S1 and Fig. S2).

Analysis on experimental Raman spectra

To show the performance of the implemented methodology in realistic environments, we applied the developed method to Raman spectra from art works. In particular, some of the experimental Raman spectra used in this research were kindly provided by Nadim

C. Scherrer from the Bern University of Applied Sciences. The experimental Raman spectra measured by the authors used in this work were acquired from private collections, and using the portable Raman equipment iHR320 with a lens of 4.5x focus (HORIBA Jobin-Yvon). The optical source was a He-Ne laser (632.8 nm) providing approximately 17 mW. The laser light was guided to the optical head by an optical fiber and directed to the samples. The same optical head collected the scattered light filtered by an edge filter. Then, it was guided to the monochromator by another optical fiber and detected by a thermoelectrically cooled CCD.

Fig. 3 presents some real-case examples of experimental Raman spectra measured from works of art, for which the proposed noise filtering technique was applied. These Raman spectra were acquired from different art works and therefore they show different shot noise realisations and different shapes of fluorescence's baseline. Specifically, the Raman spectrum before (in black) and after applying the proposed noise filtering methodology (in grey) are shown in all pictures. As it can be seen, in all the examples the Raman band shapes and positions were unchanged, and also their intensity ratios were maintained while reducing the shot noise and rejecting the baseline. Table 2 shows a comparative on the experimental Raman spectra presented in Fig. 3 carried out in the same way as for the simulated spectra. The here proposed filtering approach provided the highest signal-to-noise ratio compared to the combination of conventional denoising techniques. An example using an experimental Raman spectrum from a sample of a phthalocyanine blue pigment which shows very weak Raman bands is presented in Fig. 4. As it can be seen, the Raman bands were visibly enhanced in the denoised spectrum.

These figures provide a qualitative visual inspection of the performance of the noise filtering methodology presented in the current work. Additional figures can be found in the supplementary material (Fig. S3-S16). The denoising method reduced the influence of shot noise and removed the fluorescence's baseline without changing the shapes or positions of the Raman bands, maintaining their intensity ratios. The results show the effectiveness of the proposed denoising methodology as a fully-automated tool, that is, without requiring any user input, to help the analyst in the interpretation of Raman spectra.

Conclusion

The presence of noise is an intrinsic contribution to the difficulties in the materials analysis by Raman spectroscopy. In the case of artistic pigment analysis, external agents such as pollutants or binding media among others may increase the noise impact, thus degrading the quality of the Raman measurements. Consequently, we have developed a fully-automated denoising methodology which enhances the Raman information helping in the interpretation of the Raman spectra.

The proposed noise filtering approach uses the same novel and simple scheme for both shot noise reduction and fluorescence's baseline rejection. The method yields satisfactory results when applied to both simulated and experimental Raman spectra, providing an improved signal-to-noise ratio. Specifically, the presented noise filtering methodology does not modify the shapes of the Raman bands and maintains their intensity ratios, and therefore it reduces the interferences coming from noise sources whilst enhancing the Raman information. The proposed denoising approach is based on mathematical morphology, which retrieves the morphology of the Raman information, and therefore does not require peak recognition or previous knowledge on the baseline shape.

A benchmark study using simulated Raman spectra was presented providing a performance evaluation and comparison of the noise filtering algorithm developed in the current work and conventional denoising algorithms in common use. The results show that the presented denoising approach outperformed all other algorithms that were tested in both shot noise and baseline tests. The tests performed using experimental Raman spectra provided reliable and suitable results as well. These successful results were obtained despite of requiring no user intervention, as opposite to the other denoising techniques under test, which required some sort of user input at some point of the filtering process.

As shown by the consistency of the results, the presented noise filtering methodology has great potential as an accurate fully-automated practical method to help in the interpretation of Raman spectra, not only for artistic pigment analysis, but essentially for any material group as well.

Acknowledgements

The authors are very grateful to Nadim C. Scherrer from the Bern University of Applied Sciences for kindly providing a subset of the Raman spectra that were used in the current work.

References

[1] M. J. Pelletier, *Analytical applications of Raman Spectroscopy*, Blackwell Science, Oxford, **1999**.
[2] P.A. Mosier-Boss, S.H. Lieberman and R. Newbery, *App. Spectrosc.* **1995**; *49*, 630.
[3] Ch. Xie, Y. Li, *J. App. Physics* **2003**; *93*, 2982.
[4] M. J. Soneira, R. Pérez-Pueyo, S. Ruiz-Moreno, *J. Raman Spectrosc*. **2002** ; *33* , 599.
[5] C. J. Rowlands, S. R. Elliott, *J. Raman Spectrosc*. **2011** ; *42* , 370.
[6] C. J. Rowlands, S. R. Elliott, *J. Raman Spectrosc*. **2011** ; *42* , 1761.
[7] H. G. Schulze, et al**.**, *App. Spectrosc.* **2012** ; *66* , 757.
[8] Z. Li, D. Zhan, J. Wang, J. Huang, Q. Xu, Z. Zhang, Y. Zheng,Y. Liang, H. Wang, *Analyst*. **2013** ; *138* , 4483.
[9] I. Osticioli, A. Zoppi, W.M. Castellucci, *App. Spectrosc.* **2007**; *61*, 839.
[10] P. Matousek, M. Towrie, A. W. Parker, *J. Raman Spectrosc*. **2002**; *33*, 238.
[11] J. Zhao, M. M. Carrabba, F. S. Allen, *App. Spectrosc*. **2002**; *56*, 834.
[12] F. Rosi, et al., *J. Raman Spectrosc*. **2010**; *41*, 452.
[13] C. G. Bertinetto, T. Vuorinen, *App. Spectrosc*. **2014** ; *68* , 155.
[14] R. Pérez-Pueyo, M. J. Soneira, S. Ruiz-Moreno, *App. Spectrosc*. **2010**; *64* , 595.
[15] K. H. Liland, E. Rukke, E. Fjaervol, T. Isaksson, *Chemom. Intell. Lab. Syst.* **2011**; *109* , 51.
[16] A. Kandjani, M. J. Griffin, et al., *J. Raman Spectrosc*. **2013** ; *44* , 608.
[17] C. A. Lieber, A.Mahadevan-Jansen, *App. Spectrosc*. **2003** ; *57* , 1363.
[18] J. Zhao et al., *App. Spectrosc*. **2007** ; *61* , 1225.
[19] H. G. Schulze, et al., *App. Spectrosc*. **2011** ; *65* , 75.
[20] C. A. Lieber, A.Mahadevan-Jansen, *App. Spectrosc*. **2003** ; *57* , 1363.

[21] R. C. Carvajal, et al. *App. Spectrosc*. **2016**; *70*, 604.
[22] A. A. Urbas, S. J. Choquette, *App. Spectrosc*. **2011**; *65*, 665.
[23] A. Cao et al., *J. Raman Spectrosc*. **2007**; *38*, 1199.
[24] C. M. Galloway, E. C. Le Ru, P. G. Etchegoin, *App. Spectrosc*. **2009**; *63,* 1370.
[25] J. M. Madariaga, *J. Raman Spectrosc*. **2010** ; *41* , 1389.
[26] H. G. M. Edwards, *Spectrochim. Acta, Part A.* **2011** ; *80* , 14.
[27] H. G. M. Edwards, P. Vandenabeele, J. Jehlicka, T. J. Benoy*, Spectrochim. Acta, Part A.* **2014** ; *118* , 598.
[28] A. Tournie, L. C. Prinsloo, C. Paris, P. Colomban, B. Smith, *J. Raman Spectrosc*. **2011** ; *42* , 399.
[29] F. Casadio, A. Bezúr,I. Fiedler, K. Muir, T. Trad, S. Maccagnola, *J. Raman Spectrosc.* **2012** ; *43* , 1761.
[30] H. G. Edwards. *Historical Pigments: A Survey of Analytical Chemical Archaeometric Usage and Terminology for Forensic Art Analysis*, Encyclopedia of Analytical Chemistry, **2015**.
[31] C. Miguel, A. Claro, A. P. Gonçalves, V. S. F. Muralha, M. J. Melo, *J. Raman Spectrosc.* **2009** ; *40* , 1966.
[32] M. Irazola, M. Olivares, K. Castro, M. Maguregui, I. Martínez-Arkarazo, J. M. Madariaga, *J. Raman Spectrosc.* **2012** ; *43* , 1676.
[33] A. Domínguez-Vidal, M. J. de la Torre-López, M. J. Campos-Suñol, R. Rubio-Domene, M. J. Ayora-Cañada, *J. Raman Spectrosc.* **2014** ; *45* , 1006.
[34] S. A. Centeno, *J. Raman Spectrosc.* **2015** ; *47* , 9.
[35] G. D. Smith and R. J. H. Clark, *J. Cult. Herit*. **2002** ; *3* , 101.
[36] M. Gunn, G. Chottard, E. Rivière, J.-J. Girerd, J.-C. Chottard, *Stud. Conservat*. **2002** *; 47,* 12.
[37] A. Hernanz et al, *J. Raman Spectrosc*. **2015** ; *47* , 571.
[38] M. P. Wand, *Comput. Statist*. **2000** ; *15* , 443.
[39] D. J. Hudson, *J. Amer. Statist*. **1966** ; *61* , 1097.
[40] W. A. Fuller, Austr. J. Agric. Econ., **1969** ; *13* , 35.
[41] W. J. Studden, D. J. Van Arman, *Ann. Math. Statist.*, **1969** ; *40* , 1557.
[42] A. R. Gallant, W. A. Fuller, *J. Amer. Statist*. **1973** ; *68* , 144.
[43] S. Wold, Chemica Scripta, **1971** ; *1* , 97.
[44] P. L. Smith, *J. Amer. Statist*., **1979** ; *33* , 57.
[45] C. De Boor, *A Practical Guide to Splines*, Springer Verlag, **2001**.
[46] P. H. C. Eilers, *Anal. Chem.*, **2003** ; *75* , 3631.
[47] J. J. de Rooi, P. H. C. Eilers, *Chemom. Intell. Lab. Syst*, **2012** ; *117* , 56.
[48] L. Yang, Y. Hong, doi:10.1016/j.csda.2016.10.022
[49] P. H. C. Eilers, B. D. Marx, *Statistical Science.* **1996** ; *11* , 89.
[50] S. N. Wood, doi:10.1007/s11222-016-9666-x
[51] G. D. Knott, *Interpolating cubic splines*, Springer Science & Business Media, **2000**.
[52] C. Kooperberg, C. J. Stone, *J. Comput. And Graph. Statist*. **1992** ; *4* , 301.
[53] J. H. Friedman, B. W. Silverman, *Technometrics*, **1989** ; *1* , 3.
[54] G. Matheron, *Random Sets and Integral Geometry*, Wiley, New York, **1975**.
[55] J. Serra, *Image Analysis and Mathematical Morphology. Vol 2: Theoretical Advances*, Academic Press, New York, **1988**.
[56] E. Zhang, F. Wang, Y. Li, and X. Bai, *Bio-Med. Mater. Eng.* **2014** ; *24* , 53.
[57] V. González-Castro, J. Debayle, J. Pinoli, *Pattern Recognition Letters*. **2014** ; *47* , 50.
[58] W. K. Pratt, *IEEE Trans. on Comput.* **1972** ; *c-21* , 636.

[59] Y. Xu, J. B. Weaver, D. M. Healy Jr., J. Lu, *IEEE Transactions On Image Processing*. **1994** ; *3* , 747.

Tables:

	Linear baseline	Polynomial baseline	Sigmoidal baseline	Sinusoidal baseline
Proposed filter	**0.0319 ± 0.0097**	**0.0408 ± 0.0096**	**0.0230 ± 0.0070**	**0.0402 ± 0.0102**
PF+Wiener filter	0.0560 ± 0.0169	0.0577 ± 0.0176	0.0543 ± 0.0356	0.0768 ± 0.0326
PF+median filter	0.0575 ± 0.0140	0.0703 ± 0.0216	0.0640 ± 0.0235	0.0806 ± 0.0310
PF+wavelet filter	0.0503 ± 0.0087	0.0640 ± 0.0189	0.0565 ± 0.0205	0.0736 ± 0.0288
PF+FFT filter	0.0465 ± 0.0122	0.0498 ± 0.0151	0.0470 ± 0.0137	0.0522 ± 0.0178
PF+fuzzy filter	0.0528 ± 0.0136	0.0655 ± 0.0211	0.0593 ± 0.0227	0.0757 ± 0.0302
MF+Wiener filter	0.0350 ± 0.0066	0.0451 ± 0.0067	0.0361 ± 0.0065	0.0448 ± 0.0066
MF+median filter	0.0486 ± 0.0124	0.0486 ± 0.0124	0.0497 ± 0.0124	0.0482 ± 0.0124
MF+wavelet filter	0.0416 ± 0.0072	0.0423 ± 0.0073	0.0423 ± 0.0070	0.0417 ± 0.0072
MF+FFT filter	0.0447 ± 0.0129	0.0448 ± 0.0138	0.0460 ± 0.0119	0.0456 ± 0.0118
MF+fuzzy filter	0.0449 ± 0.0118	0.0450 ± 0.0119	0.0461 ± 0.0120	0.0447 ± 0.0128

Table 1: RMSE (mean and standard deviation) between ideal and filtered spectra using the proposed approach, and combinations of a baseline filter (conventional best-degree polynomial filter, *PF*, and morphology-based filter, *MF*) with a shot noise filter (Wiener, median, wavelet, FFT and fuzzy filters), using simulated spectra with different baseline profiles (linear, polynomial, sigmoidal and sinusoidal)

	Spectrum a	Spectrum b	Spectrum c	Spectrum d	Spectrum e	Spectrum f
Proposed filter	**28.1873**	**22.4613**	**32.3791**	**21.2516**	**31.8763**	**39.4224**
PF+Wiener filter	16.6748	11.8768	8.9701	17.4773	27.3828	23.1225
PF+median filter	16.6137	19.0401	9.2666	19.1921	30.6203	27.9877
PF+wavelet filter	15.6451	16.5602	9.3726	15.9537	29.8357	27.1419
PF+FFT filter	11.5824	14.5842	11.7827	14.0882	12.3609	24.8403
PF+fuzzy filter	16.7221	21.7162	18.5150	19.7585	30.8861	29.5737
MF+Wiener filter	20.0519	9.6524	23.9501	8.5052	30.5533	24.6272
MF+median filter	21.2140	13.2364	26.8809	14.6687	30.7646	28.6767
MF+wavelet filter	21.0518	12.0132	26.8187	11.6223	30.6575	27.2624
MF+FFT filter	21.1334	14.2837	23.6079	12.2891	30.7159	28.1590
MF+fuzzy filter	22.7607	15.2989	27.5784	15.3349	31.2740	29.6025

Table 2: SNRs of the denoised experimental Raman spectra using the proposed approach, and combinations of a baseline filter (conventional best-degree polynomial filter, *PF*, and morphology-based filter, *MF*) with a shot noise filter (Wiener, median, wavelet, FFT and fuzzy filters)

Figure captions:

Figure 1: Noise filtering scheme, aimed to reduce the shot noise and to remove the fluorescence's baseline

Figure 2: a) Graphical example of the proposed noise filtering method applied to a measured Raman spectrum of sample of a PY1 pigment powder, b) zoom for Raman shifts from 740cm^{-1} to 860cm^{-1}: 1) Shot noise reduction by fitting a cubic penalized spline through the modified closing by the minimum structuring element, 2) baseline

removal by fitting a cubic penalized spline through the modified opening by the optimal structuring element. The knot sequences are represented as black diamonds for both cases

Figure 3: Examples of experimental Raman spectra measured in art works, prior (in black) and subsequent (in grey) to apply the proposed noise filtering methodology: (a) copper-phthalocyanine blue, (b) mixture of calcite and a copper-phthalocyanine blue, (c) mixture of rutile and copper-phthalocyanine green, (d) mixture of a copper-phthalocyanine blue, carbon black and rutile, (e) mixture of a PY1, a PR4 and a copper-phthalocyanine blue, and (f) copper-phthalocyanine blue

Figure 4: Example of an experimental Raman spectrum of a copper-phthalocyanine blue pigment measured in an art work, showing very weak Raman bands (a), and resulting Raman spectrum obtained through applying the proposed noise filtering approach (b). The min-max intensities normalisation was applied to enhance visualisation