Chapter 2.1

Preprocessing of hyperspectral and multispectral images

José Manuel Amigo[a,*] and Carolina Santos[b]
[a]*Professor, Ikerbasque, Basque Foundation for Science; Department of Analytical Chemistry, University of the Basque Country, Spain; Chemometrics and Analytical Technologies, Department of Food Science, University of Copenhagen, Denmark;* [b]*Department of Fundamental Chemistry, Federal University of Pernambuco, Recife, Brazil*
**Corresponding author. e-mail: jmar@life.ku.dk*

1. Why preprocessing?

Hyperspectral imaging (HSI) and multispectral imaging (MSI) are analytical techniques based in the study of the chemical and physical behavior of the reflected or scattered light coming from a specific surface. The camera (sensor), the surface, and the light source are the analytical elements involved in each measurement. And, as any analytical element, they provide a response that is composed by the relevant analytical information, spectral noise, and different artifacts [1].

Attending to the light source, the analytical information is related to the number of photons emitted arriving to the surface. This light source is also subjected to fluctuations, whether is the sun, in a remote sensing scenario, or a halogen lamp, for bench-top hyperspectral devices. In both cases, the energy emitted might vary with time. Moreover, concerning remote sensing, that energy will pass through different atmospheric conditions before arriving to the sample.

The surface measured (sample) is hit by the photons coming from the light source. The chemical nature (composition) and the physical nature (relieve, roughness) of the surface make the photons behave differently in different parts of the sample. Thus, the number of reflected photons and their remaining amount of energy is composed not only by the analytical relevant information but also by the physical influence of the surface.

When the photons arrive to the camera in reflectance mode, they are detected by sensors that are also subjected to instrumental noise. Moreover, different sensors have different sensitivity to the photons, making the spectral

Hyperspectral Imaging. https://doi.org/10.1016/B978-0-444-63977-6.00003-1

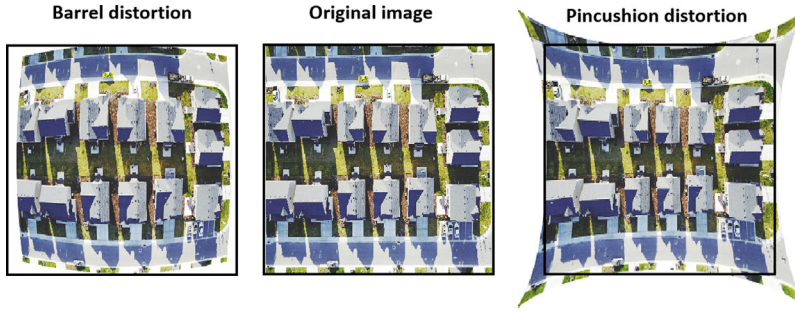

FIGURE 1 Original and distorted aerial images of houses. *The image is a free access photo taken by Blake Wheeler from Unplash website (https://unsplash.com/).*

signature dependent not only of the arriving photons but also of the quality of the sensing device.

The effects in each one of these three parts of the measurement cannot be considered as individuals, since the signal collected is a combination/mixture of all the effects. They are reflected in two main types of distortions: the distortions reflected in the geometry of the sample and the ones reflected in the spectral signal.

When the imaging area is small enough, the lens and/or Earth curvature does not cause significant distortions. However, if this is not the case, distortions such as the ones shown in Fig. 1 are observed. The figure shows an image that has been geometrically distorted due to the sensor and its position in the moment of image acquisition. On the other hand, when the field of view is small, different types of distortions can be observed, such as the ones indicated by Fig. 2 [3]. This image was taken using a bench-top instrument. This point is important to consider, since, as it was said before, the stability of light source, the sample, and the sensor plays also a fundamental role, in such a way that bench-top instruments (HSI and MSI cameras adapted to a platform in the laboratory) are more affected by certain types of artifacts than mobile/portable cameras (cameras implanted in satellites, drones, or industrial setups), and vice versa.

1.1 Spatial/geometric distortions

Spatial issues are those that arise from the geometry of the sample, from uncontrolled movements of the camera, and from the optics of the camera. The uncontrolled movements of the imaging system are especially relevant in mobile and portable cameras. In satellite imaging, for instance, the earth rotation and curvature will generate a known distortion in the acquired image; while cameras implemented in drones or airborne cameras will suffer from wind exposure and also ability of the operator during the fly. Normal geometric aberrations are optical distortions, such as the well-known pincushion and barrel distortions shown in Fig. 1, the aspect ratio between scales in the vertical direction with respect to the horizontal or viewing geometry, among others [4].

Cameras implemented in industrial setups working in conveyor belts could have deformities in the acquired image. In bench-top instruments, the platform

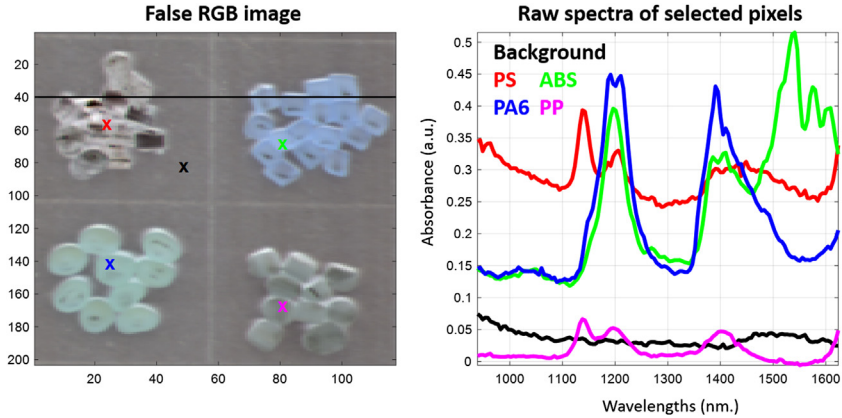

FIGURE 2 Left, false RGB of a hyperspectral image showing four different clusters of plastic pellets. The spectral range was from 940 to 1640 nm with a spectral resolution of 4.85 nm. Further information about the acquisition, type of hyperspectral camera, and calibration can be found in Ref. [3]. Right, raw spectra of the selected pixels (marked with an "x" in the left figure). The black line in the left figure indicates a missing scan line. *This figure is partially inspired by J.M. Amigo, H. Babamoradi, S. Elcoroaristizabal, Hyperspectral image analysis. A tutorial, Analytica Chimica Acta 896 (2015) 34–51. https://doi.org/10.1016/j.aca.2015.09.030.*

where the camera is implemented is normally quite robust, avoiding any problem of deformity of the acquired image. However, if confocal equipment is being employed, a specific portion of the sample could be out of focus, providing spatial distortions. Nevertheless, all cameras suffer from the roughness of the sample or irregularities of the measured surface.

One problem that is more important in bench-top instruments is the fact that many samples are not squared or are smaller than the field of view. That is, part of the acquired image contains irrelevant information of the surface where the sample is placed for the measurement. As example, Fig. 2 shows four clusters of plastics laying over a neutral black paper.

1.2 Spectral/radiometric distortions

Spectral distortions are mainly due to data recording instruments, fluctuations of the light source, and the nature of the sample. Regardless the spectral technique employed for image acquisition, the most common spectral issue is the noise. Sensors nowadays have the ability of measuring the spectral information with a high signal-to-noise ratio (SNR). Nevertheless, spectral noise will still be present (e.g., spectra shown in Fig. 2 are somehow affected by noise).

In image analysis scenario, another important aspect to consider is the saturation of light that some pixels can exhibit. Since samples are normally a distribution of different elements in the surface, it is normal, that due to the shape of the elements, their different chemical nature, and the incident angle of

the light, many pixels contain saturated information. This can be observed in the bottom right part of the false RGB image of Fig. 2, where many pixels are white due to the saturation of light in the detector.

Light scattering is also a major concern when talking about reflectance mode and, especially, about near-infrared radiation. Scattering (in additive or multiplicative way) will make the absorbance to drift the baseline of the spectra in a nonparametric manner to higher or lower values of absorbances. Scattering is due to the nature of the elements but also to the different shape of the elements, the roughness of the surface, and the ability of the radiation to penetrate in the element. See, for instance, how the baseline of the spectra shown in Fig. 2 is different for all of them. In remote sensing, the effect of the atmospheric scattering is of a major concern. The atmospheric scattering is caused by gases such as oxygen, nitrogen, ozone and also aerosols-like airborne particulate matter or clouds. When talking about Raman spectroscopy, scattered light is the source of relevant information, and fluorescence is the main cause of baseline drift, which is usually highly intense. This issue rises either from the sample itself or the highly energetic particles hitting the detector [5].

The background also plays an important role in the spectral signatures. If the background is not sufficiently neutral in the signal, it will influence the pixels belonging to the edge between the background and the elements. Background information is not of major concern on remote sensing. The signal is, however, affected by the atmosphere and the presence of aerosols and particles in suspension, which is not that relevant in bench-top instruments. For both image scenarios, one crucial spectral issue is the presence of dead pixels. Images are usually a set of a high number of pixels, and eventually, there are pixels that contain spiked signals in the spectrum, or parts of the spectrum are saturated, or simply they do not contain any information (black line in the false RGB figure of Fig. 2). This is normally generated by the malfunction of the sensor (in the case of spiked signals or no information) or by a wrong light exposure (saturated levels of light).

Different approaches have been proposed to retain the analytical signal and minimize the effect of the different issues commented before. This chapter offers a revision of the main methodologies for HSI and MSI preprocessing. Moreover, we will emphasize the benefits and drawbacks arising in the application of each one of them.

2. Geometric corrections of distortions coming from the instrument

A geometric correction is a transformation of HSI or MSI image applied to each individual channel in such a way that the distorted image is translocated to a standard reference axis (e.g., projected to coordinates in maps). This type of corrections is common in remote sensing scenarios, where a reference map of the scanned land is normally available [6–8]. The correction steps are two:

finding a proper set of coordinates in the images and in the reference map and then a step of interpolation/resampling of the distorted image to the correct reference points. The most common geometric correction method is based on ground control points (GCPs). The GCPs are reference points that are common to the distorted image and the reference map. They are normally permanent elements like road intersections or airport runways. Once the GPCs are chosen, a step of interpolation is performed. The most common resampling methods are the nearest neighbor, the bilinear, and the bicubic interpolations (Fig. 3).

The nearest neighbor resampling is the most straightforward way of resampling. It consists of assigning the nearest pixel value to the corrected pixel (Fig. 3). Bilinear interpolation, instead, considers the closest 2×2 neighborhood of known pixel values of the distorted image surrounding the unknown pixel. This method gives much smoother looking images than the nearest neighborhood (Fig. 3). Finally, the bicubic interpolation considers the 4×4 neighborhood, instead. Bicubic interpolation normally produces sharper images than the previous two methods.

3. Dead pixels, spikes, and missing scan lines

Dead pixels, spiked points in the spectra, and missing scan lines are caused by a punctual malfunction of the detector or by a saturation of light in certain points of the surface measured [9]. They can be encoded as missing values,

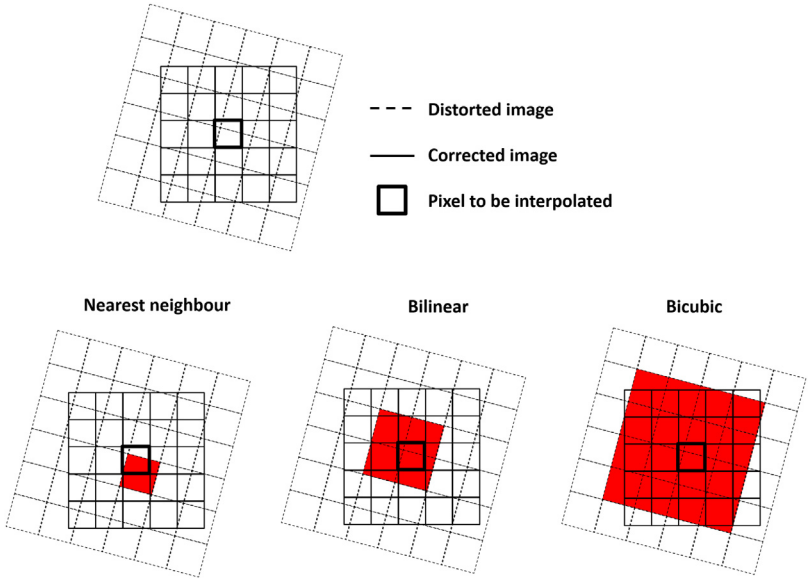

FIGURE 3 Top, distorted image (dashed squares) and corrected image (solid squares). The pixel of interest is highlighted in bold. Bottom, three of the methodologies for pixel interpolation, highlighting in each ones the pixels of the distorted image used for the interpolation.

zero or infinite values; and their location and spread might vary depending on the quality of the detector and the reflectance of the sample. They all can distort the further processing of the HSI or MSI image. This is highly problematic, since many of the routines for data mining can only handle a limited amount of missing values.

Missing scan lines and dead pixels occur when a detector fails to operate during a scan (Fig. 2). Detecting missing scan lines and dead pixels can be done using different algorithms with major or minor complexity (like thresholding techniques [10], genetic or evolutionary algorithms [11–13], or minimum volume ellipsoid (MVE) [14]). A much simpler methodology is the establishment of a predefined threshold in the number of zero values allowed in the spectrum of the pixel. Evidently, the threshold must be selected considering the nature of the data.

Spiked wavelengths are sudden and sharp rises and falls in the spectrum (Fig. 4) [15]. They are caused by an abnormal behavior of the detector or saturation of light in certain spectral region. Spikes can be normally distinguished in an easy manner. They trend to present a high deviation from the mean value of the spectrum (Fig. 4). Therefore, if a proper threshold of mean and standard deviation is chosen, they can be easily detecting, always considering difference between the normal signal, the spiked signal, and the SNR [16].

Once the missing scan lines, dead pixels, and spikes have been detected, they must be replaced by a more appropriated value. In the case of spiked wavelengths, the most straightforward way of replacing the spiked value is the substitution of that value by the mean or median of a spectral window centered in the spiked point (Fig. 4). For missing scan lines and dead pixels, many alternatives have been proposed in the literature [16–21]. Nevertheless, profiting the rich amount of spatial information that HSI and MSI normally provide, one straightforward manner to replace missing scan lines and dead pixels is substitute them by the mean or the median of the spectral of the neighboring pixels.

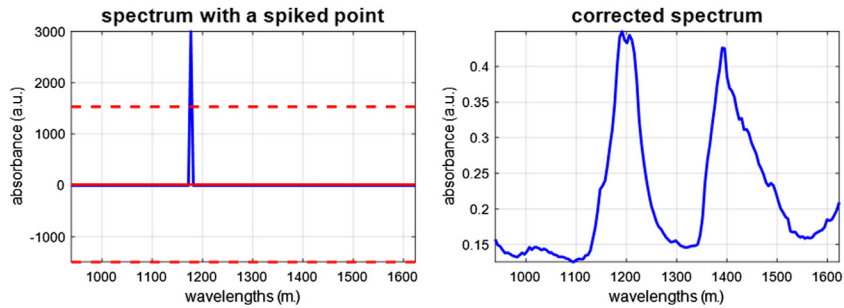

FIGURE 4 Left, spectrum containing one spiked point. The continuous red line denotes the mean of the spectrum. The upper and lower dashed red lines denote the mean ± six times the standard deviation. Right, corrected spectrum where the spike has been localized and its value substituted by an average of the neighboring spectral values.

4. Spectral preprocessing

Spectral preprocessing can be defined as the set of mathematical operations that minimize and/or eliminate the influence of undesirable phenomena affecting directly to the spectral signature obtained (e.g., light scattering, particle size effects, or morphological differences, such as surface roughness and detector artifacts [22]). In HSI, it is common to adapt the preprocessing methods coming from classical spectroscopy [5,23]. They can be divided into different families, attending their purpose. Fig. 5 shows an example of their effect in the spectral signal and in the visualization of the surface information:

- Smoothing/denoising: The instrumental noise can be partly removed by using smoothing techniques, being Savitzky−Golay the most popular one [23]. Savitzky−Golay methodologies are based on the selection of a subwindow around a specific point and calculating its projection onto a polynomial fitting of the points of the subwindow. It is simple to implement. Nevertheless, special care must be taken in the selection of the spectral subwindow, since large subwindows will eliminate informative peaks, while small windows might generate more noise.
- Scatter correction: Scattering is reflected in a drift in the baseline of the spectra (Figs. 2 and 6). That drift can be additive or multiplicative, depending on the nature of the sample and the physical interaction of the sample with the light. There are two main methods for scattering removal. The first one, standard normal variate (SNV), is the most straightforward method. It subtracts the mean of the spectrum and divides it by the standard deviation. SNV removes

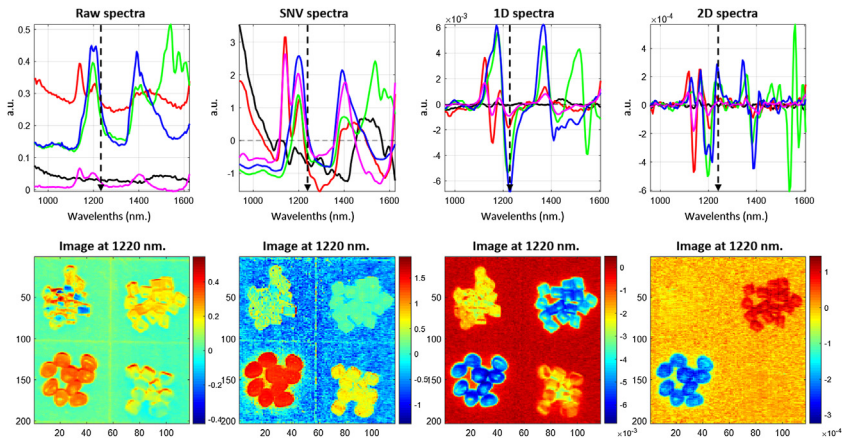

FIGURE 5 Top, raw spectra of Fig. 2 and the spectra after different spectral preprocessing methods. Bottom, the image resulting at 1220 nm for each spectral preprocessing. *This sample belongs to a data set by J.M. Amigo, H. Babamoradi, S. Elcoroaristizabal, Hyperspectral image analysis. A tutorial, Analytica Chimica Acta 896 (2015) 34−51. https://doi.org/10.1016/j.aca.2015.09.030.*

additive scattering without changing the shape of the original spectrum. Nevertheless, it cannot handle multiplicative scattering effect. Therefore, multiplicative scatter correction (MSC) is preferred when multiplicative scattering appears. MSC is also a quite straightforward method, since it projects the spectrum of pixels against one reference spectrum and then the corrected spectrum is the subtraction of the offset from the original spectrum divided by the slope [23]. The main drawback of MSC is that it depends of the correct selection of a reference spectrum. This is quite complicated to achieve in hyperspectral images, since the samples trend to be a mixture of different compounds with different spectra.

- Derivatives: The Savitzky−Golay methodology can also be used for calculating the derivative profile of the spectrum. In that sense, the subwindow of points chosen is first fitted to a polynomial degree and then the derivative is calculated. First (1D) and second (2D) derivatives are the most common ones in spectroscopy. First derivative removes the additive scattering; while second derivative removes multiplicative scattering [26]. Another effect of derivatives is their ability of highlighting minor spectral

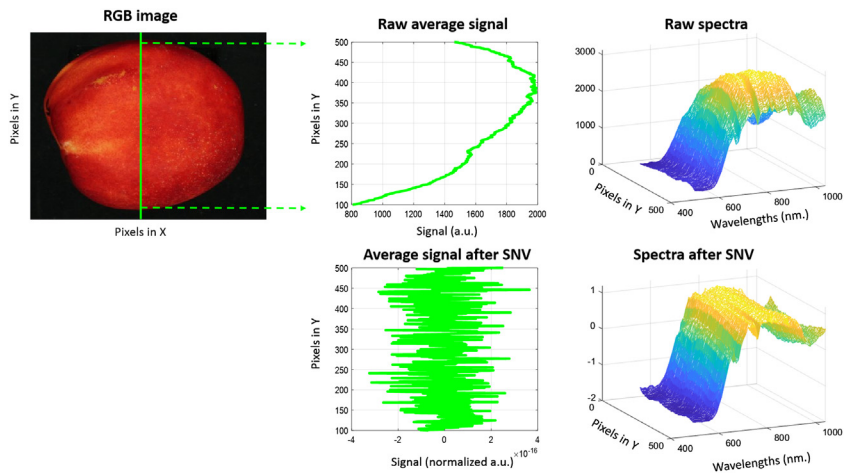

FIGURE 6 Example of spectral preprocessing for minimizing the impact of the shape of the sample in the spectra. Top left, RGB image of a nectarine. Top middle, the average signal of the pixels contained in the green line of the top left figure. Top right, spectra of the green line of the top left figure showing the effect of the curvature. Bottom middle, average signal of the preprocessed spectra (with standard normal variate (SNV)) of the green line of the top left figure. Bottom right, preprocessed spectra of the green line of the top left figure. *This sample belongs to a data set by S. Munera, J.M. Amigo, N. Aleixos, P. Talens, S. Cubero, J. Blasco, Potential of VIS-NIR hyperspectral imaging and chemometric methods to identify similar cultivars of nectarine, Food Control. 86 (2018). doi:10.1016/j.foodcont.2017.10.037; S. Munera, J.M. Amigo, J. Blasco, S. Cubero, P. Talens, N. Aleixos, Ripeness monitoring of two cultivars of nectarine using VIS-NIR hyperspectral reflectance imaging, Journal of Food Engineering 214 (2017) 29−39. https://doi.org/10.1016/j.jfoodeng.2017.06.031.*

differences. Nevertheless, special care must be taken in the choice of the derivative degree and in the subwindow size, since a high derivative degree with a small window size can create high amount of noise, while large subwindows can eliminate informative parts of the spectra [23].

One of the main properties of some spectral preprocessing methods is that they are also able to remove physical artifacts reflected in the signal. This is the case of the scattering promoted by the roughness of the sample measured or its nonplanar shape. As a matter of example, Fig. 6 shows how SNV can minimize the impact of the round shape of a nectarine [24,25]. In the figure, it is shown the raw data and the preprocessed data of one line of the hyperspectral cube. This line of spectra is clearly affected by the shape of the nectarine, effect that is minimized when SNV is applied.

5. Background removal

The selection of the regions of interest (RoI) of a sample is an important step before the analysis of the sample. This is especially relevant when the geometry of the sample does not cover the whole area measured (as illustrated in the example of Fig. 2). If the sample does not cover all the scanned area, the area left outside the sample is usually composed by highly noisy spectra, and thus, it might hamper the good performance of further models. Moreover, removing it implies a substantial saving of computing time.

As a matter of fact, successful background removal starts at image acquisition step. For bench-top instrument, it is usually possible to choose an appropriate background which facilitates its segmentation from the elements of interest. Manual selection of the RoIs, the use of specific thresholds in the histogram of the image obtained at specific wavelengths, K-means clustering, or even using the score surfaces of a principal component analysis (PCA) model are some of the methodologies that can be employed [27]. All of them have their own implication in the final result also considering the nature and the shapes of the samples in the surface. For example, Fig. 7 shows the performance of three different methodologies for removing the background of the preprocessed hyperspectral image (missing lines removed and SNV applied) shown in Fig. 7.

As it can be seen in the figure, three different methodologies provide three different answers. In this case, discerning between the edges of the samples and the background is a hard task, since the edges are pixels that contain strongly mixed information of the spectral signature of the plastic and the background. Therefore, special care must be placed in the accidental removal of informative areas of the sample.

6. Some advices on preprocessing

Preprocessing is, probably, the main engine for a successful interpretation of our hyperspectral and multispectral images. As a matter of example, Fig. 8

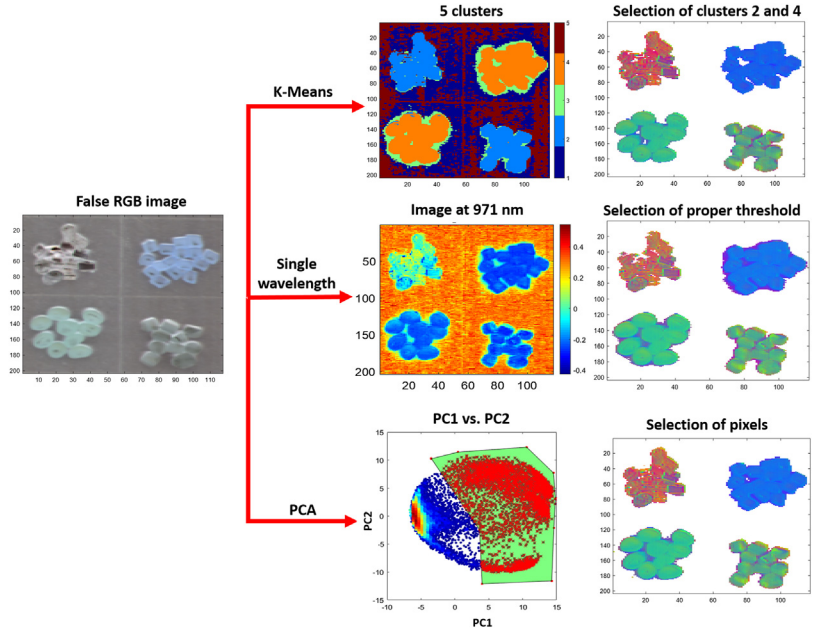

FIGURE 7 Depiction of three different methodologies for background removal. Left, false RGB image. Top, K-means analysis of the hyperspectral image in SNV and the selection of clusters 2 and 4 to create the mask. Middle, false color image obtained at 971 nm of the hyperspectral image in SNV and the result of applying a proper threshold to create the mask. Bottom, PCA scatter plot of the hyperspectral image in SNV with the selected pixels highlighted in red to create the mask. All the analysis have been made using HYPER-Tools [28], freely download in Ref. [29]. *SNV*, standard normal variate; *PCA*, principal component analysis.

shows an example of two PCA models performed on the hyperspectral image of the plastics. The first PCA model is made on the raw data (normalized prior PCA). It can be seen that even though PC1 explains almost 89% of the variance, this variance is wasted explaining the difference between the background and the plastics (Table 1).

The second PCA is made on the preprocessed data (using the spectra in first derivative and removing the background). The PCA clearly shows the spectral differences of the four plastics by using only two principal components, making the model much more understandable, parsimonious, and probably stable.

Most of the software packages for processing HSI and MSI already provide some preprocessing algorithms. And in this chapter, we have shown the benefits and drawbacks of some of the methodologies that can be applied. Nevertheless, one of the major issues in applying preprocessing methods is that the effectiveness of the correction applied must be evaluated after the application of processing algorithms. That is, as is shown in Fig. 8, the efficiency of the preprocessing methodology must be evaluated after seeing the

results of PCA. Therefore, applying preprocessing is, most of the times, a game of trial and error, although there are specific reports of genetic algorithms for preprocessing optimization [30,31].

Some main blocks have been presented here. Table 1 collects all the methods revisited here giving an account of their major benefits and drawbacks. One question that might arise is the order and number of preprocessing steps that must be used. Unfortunately, there is not a specific answer for that question. Or, better said, the answer can be again a trial error game. Sometimes the background is easily removed from the raw data and then the data included in the RoI are preprocessed; and sometimes a spectral preprocessing is needed for removing the background while another spectral preprocessing is needed for the analysis of the sample. In any case, there are some major advices that can be given:

- Parsimony. The simpler, the better: Preprocessing normally changes the spatial and the spectral information, in such a way that those changes can remove informative parts in our image. Moreover, it can also introduce artifacts or generate the loss of important information if the proper method is not selected or correctly applied. Therefore, the simpler a preprocessing methodology is, the better, as long as we achieve the desired results.
- Spatial and spectral corrections are connected: Smoothing the spectra in an HSI sample will not only remove the spectral noise but also will smooth the images arising from the data. The application of spectral corrections has an implication in the surface and vice versa.
- There is a price to pay: By applying preprocessing, there will always be lost information. It is our responsibility to lose only the information that we can consider noise and stay with the analytical relevant information.

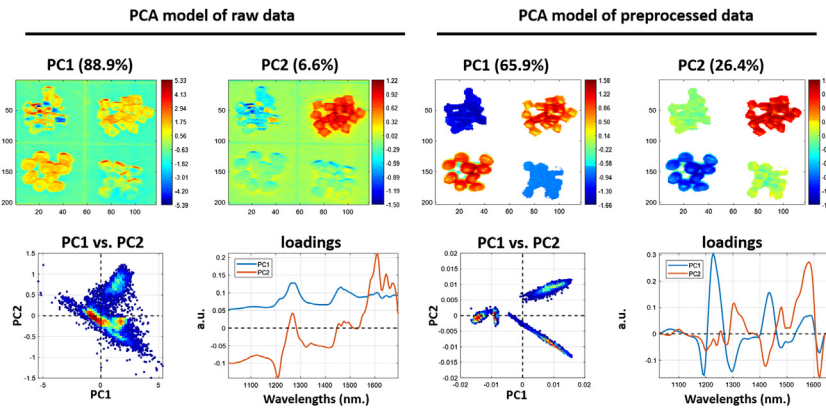

FIGURE 8 Comparison of two PCA models performed in the hyperspectral image of the plastics [3]. For each PCA model, the score surfaces of the first two PCs, the scatter plot of PC1 versus PC2 and the corresponding loadings are shown. All the analyses have been made using HYPER-Tools [28], freely download in Ref. [29]. *PCA*, principal component analysis; *PC*, principal component.

TABLE 1 Summary of the main preprocessing steps, the techniques for their application, and their benefits/drawbacks.

Preprocessing step	Techniques	Types of images	Benefits	Drawbacks
Dead pixels	Detection			
	Median spectra/ thresholding [10,32]	MSI and HSI	Easy to implement and calculate	Highly dependent of the signal-to-noise ratio. Risk of false positives.
	Genetic/ evolutionary algorithms [9,12,13]	MSI and HSI	Robust and reliable	To find the best combination of parameters to optimize the models.
	Chosen criteria	MSI and HSI	Easy to implement and calculate	Difficulties in finding the proper threshold.
	Suppression			
	Neighboring interpolation	MSI and HSI	Easy to implement and calculate	If the cluster of dead pixels is big, there could be the risk of losing resolution in this area
Spikes	Detection			
	Manual inspection	MSI and HSI	Robust and reliable	Time-consuming, specially in HSI images
	Neighbor filters [15,18,19]		Robust and reliable	To find the best combination of parameters to optimize the filters. When the background is an important part of the image, there may be problems to differentiate the spikes.

TABLE 1 Summary of the main preprocessing steps, the techniques for their application, and their benefits/drawbacks.—cont'd

Preprocessing step	Techniques	Types of images	Benefits	Drawbacks
	Wavelets [17,20,21]		Robust and reliable	The selection of the proper wavelet (in the spatial and the spectral channels) for each type of image.
	Chosen criteria		Robust and reliable	Difficulties in finding the proper threshold.
	Suppression			
	Neighbor interpolation [15,17–21]		Easy to implement and calculate	
Background/ ROI	PCA thresholding [10,33]	MSI and HSI	Robust selection of a specific area based of PC scores images	The selection of the proper threshold is tedious and not obvious in some situations.
	Manual		Selection of the desired area	Time-consuming, specially working with time series images or large data sets
	K-means [34]		Easy to implement and calculate	
Spectral preprocessing	Denoising			
	Savitzky–Golay smoothing [23]	HSI	Easy to implement	To find the best combination of parameters to optimize the filter, especially the window size.

Continued

TABLE 1 Summary of the main preprocessing steps, the techniques for their application, and their benefits/drawbacks.—cont'd

Preprocessing step	Techniques	Types of images	Benefits	Drawbacks
	Scatter correction			
	MSC, SNV [23]		SNV does not change the shape of the spectra	Sometimes the suppression of artifacts is not totally achieved. MSC and derived techniques need additional information and may change the shape of the spectra.
	Derivatives [23]		Removal of different baseline artifacts	To find the best combination of parameters to optimize the filter. Especially the window size and the derivative order.
Geometric corrections	Nearest neighbor interpolation [35]	MSI and HSI	Easy to implement	To find the proper set of reference points to make proper interpolation.
	Bilinear interpolation [35]		Easy to implement Smooth edges are created	To find the proper set of reference points to make proper interpolation.
	Bicubic interpolation [35]		Easy to implement Smooth edges are created Sharper images	To find the proper set of reference points to make proper interpolation.

HIS, Hyperspectral imaging; *MSC*, Multiplicative scatter correction; *MSI*, Multispectral imaging; *PCA*, Principal component analysis; *SNV*, Standard normal variate.
Extracted and reproduced from M. Vidal, J.M. Amigo, Pre-processing of hyperspectral images. Essential steps before image analysis, Chemometrics and Intelligent Laboratory Systems 117 (2012) 138–148. https://doi.org/10.1016/j.chemolab.2012.05.009 and modified with permission of Elsevier.

References

[1] M. Vidal, J.M. Amigo, Pre-processing of hyperspectral images. Essential steps before image analysis, Chemometrics and Intelligent Laboratory Systems 117 (2012) 138–148, https://doi.org/10.1016/j.chemolab.2012.05.009.

[2] Unplash, Unplash, (n.d.). https://unsplash.com/.

[3] J.M. Amigo, H. Babamoradi, S. Elcoroaristizabal, Hyperspectral image analysis. A tutorial, Analytica Chimica Acta 896 (2015) 34–51, https://doi.org/10.1016/j.aca.2015.09.030.

[4] P.K. Varshney, M.K. Arora, Advanced Image Processing Techniques for Remotely Sensed Hyperspectral Data, Springer, 2004.

[5] T. Bocklitz, A. Walter, K. Hartmann, P. Rösch, J. Popp, How to pre-process Raman spectra for reliable and stable models? Analytica Chimica Acta 704 (2011) 47–56, https://doi.org/10.1016/j.aca.2011.06.043.

[6] T. Toutin, Review article: geometric processing of remote sensing images: models, algorithms and methods, International Journal of Remote Sensing 25 (2004) 1893–1924, https://doi.org/10.1080/0143116031000101611.

[7] N.G. Kardoulas, A.C. Bird, A.I. Lawan, Geometric correction of SPOT and landsat Imagery : a comparison of M a p and GPS-derived control points, Photogrammetric Engineering & Remote Sensing 62 (1996) 1173–1177.

[8] A.J. De Leeuw, L.M.M. Veugen, H.T.C. Van Stokkom, Geometric correction of remotely-sensed imagery using ground control points and orthogonal polynomials, International Journal of Remote Sensing 9 (2007) 1751–1759, https://doi.org/10.1080/01431168808954975.

[9] R. Leardi, Genetic algorithms in chemometrics and chemistry: a review, Journal of Chemometrics 15 (2001) 559–569, https://doi.org/10.1002/cem.651.

[10] J. Burger, P. Geladi, Hyperspectral NIR image regression part I: calibration and correction, Journal of Chemometrics 19 (2005) 355–363. https://doi.org/10.1002/cem.986.

[11] R. Leardi, Experimental design in chemistry: a tutorial, Analytica Chimica Acta 652 (2009) 161–172. https://doi.org/10.1016/j.aca.2009.06.015.

[12] B. Walczak, Outlier detection in multivariate calibration, Chemometrics and Intelligent Laboratory Systems 28 (1995) 259–272. https://doi.org/10.1016/0169-7439(95)80062-E.

[13] P. Vankeerberghen, J. Smeyers-Verbeke, R. Leardi, C.L. Karr, D.L. Massart, Robust regression and outlier detection for non-linear models using genetic algorithms, Chemometrics and Intelligent Laboratory Systems 28 (1995) 73–87. https://doi.org/10.1016/0169-7439(95)80041-7.

[14] C. Junghwan, P.J. Gemperline, Pattern recognition analysis of near-infrared spectra by robust distance method, Journal of Chemometrics 9 (1995) 169–178. https://doi.org/10.1002/cem.1180090304.

[15] L. Zhang, M.J. Henson, A practical algorithm to remove cosmic spikes in Raman imaging data for pharmaceutical applications, Applied Spectroscopy 61 (2007) 1015–1020. https://doi.org/10.1366/000370207781745847.

[16] Z. Nenadic, J.W. Burdick, Spike detection using the continuous wavelet transform, IEEE Transactions on Biomedical Engineering 52 (2005) 74–87. https://doi.org/10.1109/TBME.2004.839800.

[17] F. Ehrentreich, L. Summchen, Spike removal and denoising of Raman spectra by wavelet transform methods, Analytical Chemistry 73 (2001) 4364−4373. https://doi.org/10.1021/ac0013756.

[18] C.J. Behrend, C.P. Tarnowski, M.D. Morris, Identification of outliers in hyperspectral Raman image data by nearest neighbor comparison, Applied Spectroscopy 56 (2002) 1458−1461. https://doi.org/10.1366/00037020260377760.

[19] C.V. Cannistraci, F.M. Montevecchi, M. Alessio, Median-modified Wiener filter provides efficient denoising, preserving spot edge and morphology in 2-DE image processing, Proteomics 9 (2009) 4908−4919. https://doi.org/10.1002/pmic.200800538.

[20] K. Koshino, H. Zuo, N. Saito, S. Suzuki, Improved spike noise removal in the scanning laser microscopic image of diamond abrasive grain using wavelet transforms, Optics Communications 239 (2004) 67−78. https://doi.org/10.1016/j.optcom.2004.05.056.

[21] P. Du, W.A. Kibbe, S.M. Lin, Improved peak detection in mass spectrum by incorporating continuous wavelet transform-based pattern matching, Bioinformatics 22 (2006) 2059−2065. https://doi.org/10.1093/bioinformatics/btl355.

[22] J.M. Amigo, Practical issues of hyperspectral imaging analysis of solid dosage forms, Analytical and Bioanalytical Chemistry 398 (2010) 93−109. https://doi.org/10.1007/s00216-010-3828-z.

[23] A. Rinnan, F. van den Berg, S.B. Engelsen, Review of the most common pre-processing techniques for near-infrared spectra, TRAC Trends in Analytical Chemistry 28 (2009) 1201−1222, isi:000272098600016.

[24] S. Munera, J.M. Amigo, N. Aleixos, P. Talens, S. Cubero, J. Blasco, Potential of VIS-NIR hyperspectral imaging and chemometric methods to identify similar cultivars of nectarine, Food Control 86 (2018). https://doi.org/10.1016/j.foodcont.2017.10.037.

[25] S. Munera, J.M. Amigo, J. Blasco, S. Cubero, P. Talens, N. Aleixos, Ripeness monitoring of two cultivars of nectarine using VIS-NIR hyperspectral reflectance imaging, Journal of Food Engineering 214 (2017) 29−39. https://doi.org/10.1016/j.jfoodeng.2017.06.031.

[26] P. Geladi, D. MacDougall, H. Martens, Correction for near-infrared reflectance spectra of meat, Applied Spectroscopy (1985). https://doi.org/10.1366/0003702854248656.

[27] N.R. Pal, S.K. Pal, A review on image segmentation techniques, Pattern Recognition 26 (1993) 1277−1294.

[28] N. Mobaraki, J.M. Amigo, HYPER-Tools. A graphical user-friendly interface for hyperspectral image analysis, Chemometrics and Intelligent Laboratory Systems 172 (2018). https://doi.org/10.1016/j.chemolab.2017.11.003.

[29] J.M. Amigo, HYPER-tools Official Website, (n.d.). Hypertools.org (accessed March 10, 2019).

[30] O. Devos, L. Duponchel, Parallel genetic algorithm co-optimization of spectral preprocessing and wavelength selection for PLS regression, Chemometrics and Intelligent Laboratory Systems 107 (2011) 50−58. https://doi.org/10.1016/j.chemolab.2011.01.008.

[31] R.M. Jarvis, R. Goodacre, Genetic algorithm optimization for pre-processing and variable selection of spectroscopic data, Bioinformatics 21 (2005) 860−868. https://doi.org/10.1093/bioinformatics/bti102.

[32] P. Geladi, J. Burger, T. Lestander, Hyperspectral imaging: calibration problems and solutions, Chemometrics and Intelligent Laboratory Systems 72 (2004) 209−217, isi:000222986900012.

[33] J. Burger, P. Geladi, Hyperspectral NIR image regression part II: dataset preprocessing diagnostics, Journal of Chemometrics 20 (2006) 106−119. https://doi.org/10.1002/cem.986.

[34] N. Dhanachandra, K. Manglem, Y.J. Chanu, Image segmentation using K -means clustering algorithm and subtractive clustering algorithm, Procedia Computer Science 54 (2015) 764–771. https://doi.org/10.1016/j.procs.2015.06.090.

[35] D. Han, Comparison of commonly used image interpolation methods, in: Proc. 2nd Int. Conf. Comput. Sci. Electron. Eng. (ICCSEE 2013), 2013, pp. 1556–1559.