Journal of Pharmaceutical and Biomedical Analysis 48 (2008) 533–553

Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

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

Review

# Pharmaceutical applications of vibrational chemical imaging and chemometrics: A review

C. Gendrin [a,b,*], Y. Roggo [a], C. Collet [b]

[a] F. Hoffmann-La Roche A.G., Basel, Switzerland
[b] LSIIT - UMR CNRS 7005, Strasbourg University, France

ARTICLE INFO

*Article history:*
Received 28 March 2008
Received in revised form 4 August 2008
Accepted 9 August 2008
Available online 22 August 2008

*Keywords:*
Near infrared
Raman
Infrared
Chemical imaging
Pharmaceutics
Spectroscopy

ABSTRACT

The emergence of chemical imaging (CI) has gifted spectroscopy an additional dimension. Chemical imaging systems complement chemical identification by acquiring spatially located spectra that enable visualization of chemical compound distributions. Such techniques are highly relevant to pharmaceutics in that the distribution of excipients and active pharmaceutical ingredient informs not only a product's behavior during manufacture but also its physical attributes (dissolution properties, stability, etc.). The rapid image acquisition made possible by the emergence of focal plane array detectors, combined with publication of the Food and Drug Administration guidelines for process analytical technology in 2001, has heightened interest in the pharmaceutical applications of CI, notably as a tool for enhancing drug quality and understanding process. Papers on the pharmaceutical applications of CI have been appearing in steadily increasing numbers since 2000. The aim of the present paper is to give an overview of infrared, near-infrared and Raman imaging in pharmaceutics. Sections 2 and 3 deal with the theory, device set-ups, mode of acquisition and processing techniques used to extract information of interest. Section 4 addresses the pharmaceutical applications.

© 2008 Elsevier B.V. All rights reserved.

**Contents**



*Abbreviations:* AH, agglomerative hierarchical; ANN, artificial neural network; API, active pharmaceutical ingredient; ASA, acetylsalicylic acid; ATM, atomic force microscopy; ATR, attenuated total reflection; BSS, blind source separation; BTEM, band target entropy minimization; CCA, cosine correlation analysis; CCD, charge coupled detectors; CI, chemical imaging; CLS, classical least squares; csiFCM, cluster size insensitive fuzzy-C mean; 2D/3D, two-/three-dimensional; DA, discriminant analysis; DAC, directed agglomeration clustering; DCLS, direct classical least squares; DESI-MS, desorption electrospray ionization linear ion trap mass spectrometry; $D_2O$, deuterium oxide; DR, diffuse reflection; EMSC, extended multiplicative signal correction; EVA, ethylene vinyl acetate; FCM, fuzzy-C mean; FNNLS, fast non-negative least squares; FPA, focal plane array; FFT, fast Fourier transform; FIR, far-infrared; FOV, field of view; FT, Fourier transform; FTIR, Fourier transform infrared; HPLC, high performance liquid chromatography; HPMC, hydroxypropylmethylcellulose; ICA, independent component analysis; IR, infrared; ITTFA, iterative target transformation factor analysis; KM, K-means; KSFA, key-set factor analysis; LCTF, liquid crystal tunable filter; LDA, linear discriminant analysis; LLS, laser light scattering; LUT, look-up table; MBCD, methyl-β-cyclodextrin; MCR-ALS, multivariate curve resolution-alternating least squares; MCT, mercury cadmium telluride; MIA, multivariate image analysis; MIR, mid-infrared; MLF-ANN, multilayer feed-forward-artificial neural network; MLP-ANN, multilayer perception-artificial neural network; MNF, maximum noise fraction; MSC, multiplicative scatter correction; NA, numerical aperture; NIR, near-infrared; NMF, non-negative matrix factorization; OLS, ordinary least squares; OPA, orthogonal projection analysis; PARAFAC:, parallel factor; PAT, process analytical technology; PBS, phosphate buffered saline; PCA, principal component analysis; PDMS, polydimethylsiloxane; PEG, polyethylene glycol; PEO, polyethylene oxide; PLGA, poly(lactic-*co*-glycolic acid); PLS, partial least squares; PMF, positive matrix factorization; RGB, red–green–blue; ROI, regions of interest; SA, salicylic acid; S.D., standard deviation; SIM, spectral identity mapping; SIMCA, soft independent modeling of class analogies; SIMPLISMA, simple-to-use interactive self-modeling mixture analysis; SMCR, self-modeling curve resolution; SNR, signal to noise ratio; SNV, standard normal variate; SVM, support vector machine; UV, ultraviolet; WEFA, window evolving factor analysis.

* Corresponding author at: F. Hoffmann-La Roche A.G., Grenzacherstrasse, Basel, Switzerland.
*E-mail address:* christelle.gendrin@roche.com (C. Gendrin).

0731-7085/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpba.2008.08.014

## 1. Introduction

Vibrational spectroscopy encompasses near-infrared (NIR), mid-infrared (MIR) and Raman spectroscopy. Aided by relentless instrumentational advance, these complementary techniques have now entered common use for the study of solid-state samples. They allow both quantitative and qualitative analysis and can also be deployed in-line. They have therefore found many applications in the industrial pharmaceutical setting [1,2].

In 1949, in *Nature*, Barer and Cole reported the acquisition of spatially resolved spectra using a microscope [3]. This publication opened up new opportunities for spectroscopy by allowing microscopic samples to be analyzed. Four decades later, in 1988, Harthcock and Atkin obtained the first chemical map in the MIR range [4] using a microscope and moving stage. By revealing information that is both spectral and spatial, the technique can identify and localize compounds. Subsequent refinements have been exponential. Development of the first microscope-mounted focal plan array (FPA) detectors increased enthusiasm for chemical imaging (CI) [5]. Fast and robust acquisition is now possible in the NIR and MIR ranges and also with Raman spectroscopy. Almost all chemical compounds in a sample can be visualized within minutes. Applications have since increased in various fields, from waste sorting [6] to biological tissue [7] and food quality [8]. More recently, the introduction in 2001 of the concepts of process analytical technology (PAT) [9] and quality by design sparked fresh interest in the pharmaceutical industry. PAT is a generic term for monitoring quality and performance parameters in production-based systems. It has two objectives: (1) it requires the product to be monitored not only at the end of the production line, as at present, but also all along the line, in order to enhance quality; (2) it encourages the development of new process-monitoring methods and systems, thereby enhancing understanding of the process itself. One undeniable factor influencing the physical attributes of solid dosage forms is compound distribution. For example, heterogeneous compound distribution can decrease the rate of tablet dissolution [10] or prompt process troubleshooting over poor powder flow or sticky tablets [11]. CI is an ideal tool for resolving these issues when spatial information is required [12]. Unsurprisingly, publications dealing with hyperspectral imaging have multiplied since 2000.

CI has also raised new data-processing challenges. A single acquisition may record thousands of images across numerous wavelengths. The resulting image stack forms a three-dimensional (3D) matrix, or data cube, spanning two spatial dimensions with a series of wavelengths making up the third (spectral) axis (Fig. 1). There are two challenges: (1) the data cube may be viewed as spatially located spectra, with the processing tools of classical spectroscopy being applied to single spectra; (2) the data may be viewed as images, with image-processing tools being used to extract higher-quality spatial information. CI thus combines the techniques of spectroscopy and signal and image processing, making it a truly multidisciplinary discipline.

This review will not only offer a comprehensive presentation of the pharmaceutical applications of vibrational imaging but also cover and discuss the instrumentation and data-processing techniques available. Section 2 (vibrational spectroscopy theory and instrumentation) encompasses several important aspects of imaging, such as acquisition time, calibration, and resolution (both spectral and spatial). Section 3 presents the processing techniques available for extracting relevant information from data cubes. Both parts refer to publications in other fields, such as agriculture and biology, to corroborate and illustrate the

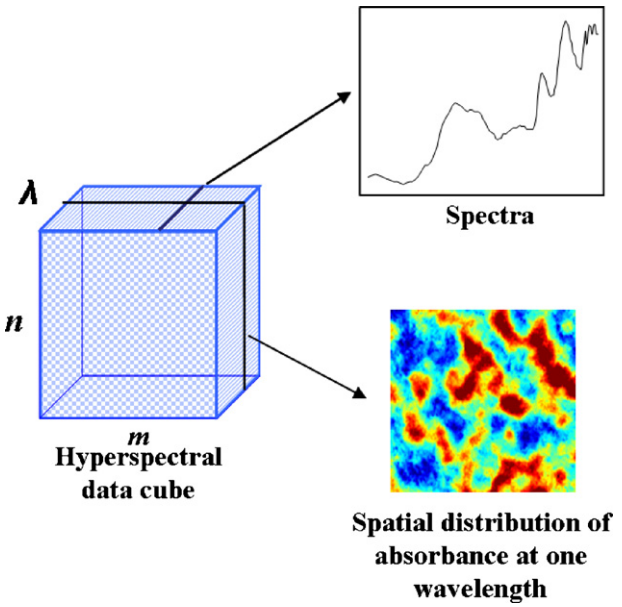


**Fig. 1.** Three-dimensional data cube recorded during a chemical imaging experiment.

points made. Section 4 is devoted exclusively to pharmaceutical applications.

## 2. Basic principles of vibrational spectroscopy and imaging

### 2.1. Vibrational spectroscopy

As its name indicates, vibrational spectroscopy uses the particular spectral range from the electromagnetic spectrum (13,000–10 $cm^{-1}$ or 0.76–1000 μm, depending on the units employed) at which the chemical bonds in molecules vibrate. This spectral region encompasses three subdivisions: the far-infrared (FIR: 400–10 $cm^{-1}$ or 26–1000 μm), MIR (4000–400 $cm^{-1}$ or 2.6–26 μm), and NIR (13,000–4000 $cm^{-1}$ or 0.76–2.6 μm), named in relation to the visible region. Two units are used in vibrational spectroscopy: $cm^{-1}$ (wavenumbers) or nm. The choice of one of the units depends either on the type of spectrometer (dispersive vs. Fourier transform (FT)) or to avoid too large numbers in the NIR range where nm is more often used. The relationship between the two units is given by

$$[cm^{-1}] = \frac{1}{[nm] \times 10^{-7}} \quad (1)$$

In NIR and MIR spectroscopy sample absorbance is recorded at each wavelength. NIR spectra arise from recording molecular overtone and combination vibrations. The MIR spectrum records the absorbance of light (energy) at the vibrational and rotational frequencies of the atoms within the molecule [13]. These frequencies can be simplistically described by a spring model in which the atoms of a molecule shrink and elongate around an equilibrium position. Fig. 2a shows the MIR spectrum of an active pharmaceutical ingredient (API). In this spectrum, the absorbance peaks are numerous and fine due to these fundamental transitions, allowing the identification of the functional group and molecule. On the other hand, absorbance is strong. Consequently, either the samples must be diluted into nonabsorbing matrix to avoid saturating the detector or advanced sampling techniques such as attenuated total reflection (ATR) must be used [14] (Section 2.4), making mid-IR impractical for in-line analysis of solids.

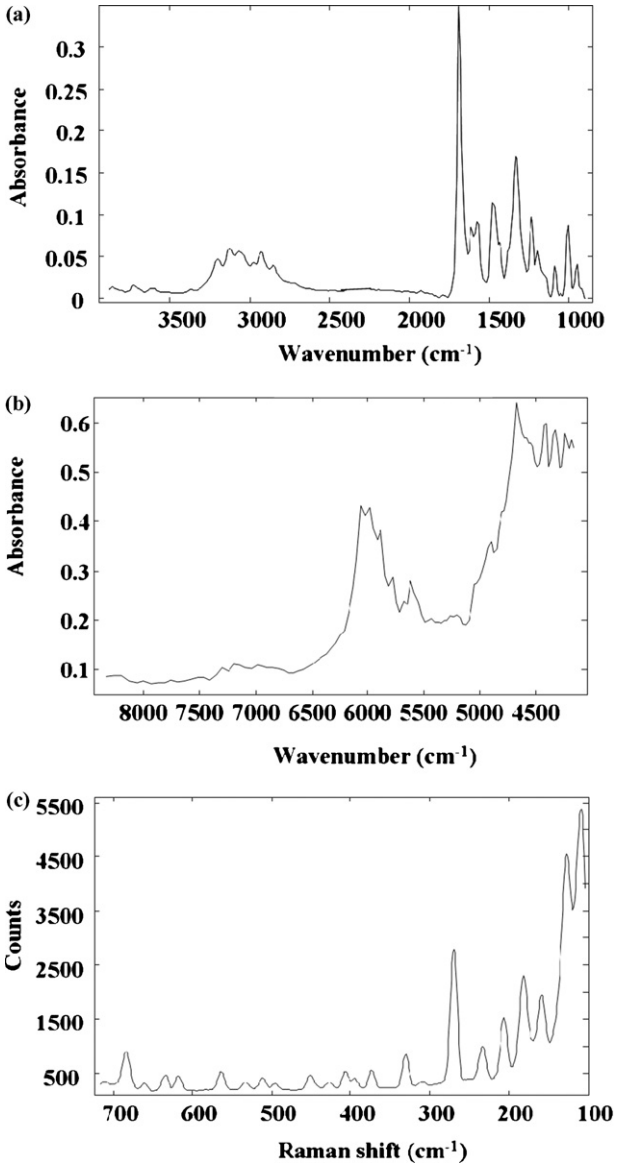


**Fig. 2.** Mid-infrared (a), near-infrared (b), and Raman (c) spectra of an API.

Combination and overtone vibrations are recorded in the NIR range. In the NIR spectrum of an API (Fig. 2b), the spectral range is narrower than in the MIR range. Peaks are also broader and overlap, complicating the clear identification of chemicals from their peak positions. Statistical analysis of the spectra is often required for molecule identification and quantification. However, since the absorbance signal is weak, due to low probability of occurrence of overtone and combination vibrations, samples require no preparation and analysis is nondestructive. The samples might be analyzed through glass because it is transparent in the NIR due to a lack of significant X–H bonds and also through plastic depending on the spectral window. The development of computer resources, combined with user-friendlier devices (detectors that do not need to be frozen), makes NIR the spectroscopy of choice for in-line applications.

Raman spectroscopy [15] involves not the absorption of light but its scattering. Samples are illuminated by a monochromatic beam, usually a laser. When the beam hits the molecules, most of the light is scattered at the same wavelength (elastic Rayleigh scattering). However, an infinitesimal part (1 per $10^9$ up to 1 per

$10^6$) is scattered at shorter or longer wavelengths. This difference, or wavelength shift, can be related to vibrations and thus chemical bonds. The number of diffused photons at wavelength shifts is recorded during a Raman experiment. The Raman spectrum of an API (Fig. 2c) shows many narrow peaks across a wide spectral range and clearly identifies the molecules. It complements MIR spectroscopy because the selection rules are not the same in the two spectroscopic techniques. For example, S–S and C=C bands are weak in the IR but among the strongest in the Raman spectrum [15,16]. The advantage of Raman spectroscopy is that it can analyze the sample through glass or plastic due to the high energy of the incident beam. Samples also need minimum preparation. However, fluorescence (emission of photons due to electronic transitions of atoms when exposed to an electromagnetic radiation) may prevent detection of the Raman bands.

### 2.2. Device set-up for microspectroscopy

Microspectroscopy systems consist of four main parts [17]:

(1) *Light source*. A single polychromatic thermal source is generally used for MIR and NIR spectroscopy. An inert solid electrically heated to 1500–2200 K irradiates uniformly in the IR spectral range. Silicon-carbide is used in MIR and a tungsten filament in NIR. In Raman spectroscopy, lasers emit the excitation beam. Typical wavelengths of excitation in the visible range are 488 nm (blue), 514 nm (green), 632 nm (red), and in the NIR 785 nm, 830 nm and 1064 nm. The choice of wavelength will depend on the solid-state properties of the sample and its fluorescence. Longer excitation wavelengths prevent fluorescence but decrease also the intensity of Raman band roughly proportional to the fourth power of the wavelength [16].
(2) *Splitter*. The various techniques that have been developed differ in their advantages and disadvantages. FT, tunable filter, and diffraction grating spectrometers are the three main types used in vibrational imaging:
   - FT spectrometers record information from several wavelengths simultaneously. Their theoretical basis is described elsewhere [18]. Their many advantages include: rapid (typically ~1 s) acquisition across the whole spectral range; high spectral resolution, down to 2 $cm^{-1}$; high energy available to the detector (because no slits are used); and high wavelength repeatability thanks to using a laser for reference spectra. An FT interferometer can be used for both MIR and NIR spectroscopy [19] but also Raman spectroscopy [20].
   - Filters are useful for focusing on specific wavelengths. They also dispense with moving parts in the spectrometer [21]. Several filters can be mounted on a wheel to select several wavelengths. An alternative is a tunable filter [22], which electronically controls spectral transmission by applying a voltage. The liquid crystal tunable filter (LCTF) has become the most popular technique for global imaging. Its advantage over the filter wheel is that it can record more than 100 images at different wavelengths. Filters are mainly used in NIR and Raman hyperspectral imaging.
   - A diffraction grating has a large number of parallel lines or slits separated by a distance comparable to the wavelength of light. When a polychromatic ray of light hits the grating, it is dispersed in several directions and the angle of diffraction is dependent on the wavelength. With single-point detectors only one spectral point can be acquired per position of the prism and detector. The full spectrum can be acquired by rotating either part stepwise. Line detectors enable several wavelengths to be acquired simultaneously [23–25]. Slits at the entrance and exit of the gratings remove parasitic light. However, narrow slits can reduce drastically the amount of signal reaching the detector whereas large slits might decrease the spectral resolution of the spectrometer. High detector sensitivity and high source intensity in the NIR range make it suitable for NIR [24] and also Raman spectroscopy.
(3) *Detector*. Photon detectors are the most widely used in MIR and NIR spectroscopy to record the signal after wavelength separation. In NIR, lead sulfide (PbS), indium antimonide (InSb), and uncooled indium gallium arsenide (InGaAs) are commonly used, whereas in MIR mercury cadmium telluride (HgCdTe or MCT) is most used due to wide spectral sensitivity (from 2 μm to 20 μm, depending on the Hg/Cd concentration ratio) [5,19,26–28]. Raman detectors are mostly 2D charge coupled detectors (CCDs) [29].
(4) *Optics*. The microscope is fitted with optical elements for selecting spatial resolution. The choice of magnification levels depends on the study aim. Typically 6×, 15×, and 32× objectives are used on a mid-IR or NIR microscope [30], and 50× and 100× objectives on a Raman microscope. Additionally, Raman microscopes are often mounted with a confocal aperture to tune the penetration depth of the beam and thereby analyze only few micrometers of the sample. Such systems can also focus the beam at different sample depths to determine subsurface composition [31].

### 2.3. From microspectroscopy to imaging

#### 2.3.1. Acquisition of a hyperspectral cube

In classical spectroscopy a spectrum reflects the integrated spectral information of the sample surface which depends on the spot size generated by the beam of light. Beginning in the early 1990s, CI systems became available for acquiring spectral and spatial information simultaneously. The spatial locations of the spectra identify chemical species inside the samples and map their distributions.

Three techniques are available for generating a hyperspectral cube (Fig. 3):

(1) *Point mapping* (Fig. 3a). This technique, the most used up to 2000, consists of a classic spectrometer combined with a moving stage [4]. The user defines a regular grid of spatial positions above the sample surface. A spectrum is measured at one position, the sample moves to the next measurement point on the grid, a further spectrum is recorded, and so on for all positions in the area defining the image. Almost all manufacturers now offer point mapping for microscopic spectroscopy.
(2) *Line imaging* (Fig. 3b). The detector acquires the spatial and spectral dimensions simultaneously. As with point mapping, the system acquires spectra according to predefined spatial positions and the line is moved right to left and up to down to cover the whole area. For example, a line acquiring 21 and 16 spectra simultaneously may be used for Raman [32] and Fourier transform infrared (FTIR) [33] imaging, respectively.
(3) *Focal plane array* (FPA) optical detectors (Fig. 3c) [5,27] are composed of several thousand elements forming a matrix of pixels. They enable thousands of spectra to be acquired simultaneously [34]. Such systems have become increasingly popular in the last 10 years and several systems are now commercially available.

#### 2.3.2. Comparison of acquisition processes

Few papers in the literature compare point mapping, line mapping and widefield imaging. The strongest argument in favor of widefield imaging is the faster acquisition time, making it an

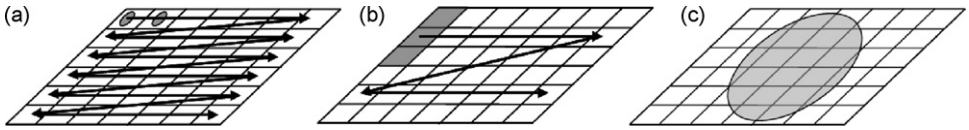


**Fig. 3.** Three methods of acquiring a hyperspectral imaging data cube: (a) point mapping, (b) line mapping, and (c) widefield approach.

instrument of choice when analyzing dozens of samples. A study comparing acquisition times between point mapping, line mapping and global Raman imaging systems for the same number of pixels and same illumination time found that an experiment using point mapping took about 12 days with 4 s of acquisition time/pixels, line scanning about 30 min with 4 s of acquisition time/line, and widefield about 32 s for the acquisition of one image at one wavelength [34]. Another advantage of widefield is the ability to shorten acquisition time and reduce data volume by choosing a specific wavelength, but this presupposes detailed foreknowledge of specific chemical vibrations. In another comparison of acquisition times using point mapping and widefield, 12.5 h were needed to acquire 2500 spectra over a 1 $mm^2$ area with a spatial resolution of 20 $\mu m \times 20\ \mu m$ and spectral resolution of 16 $cm^{-1}$ using an FT-NIR spectrometer versus 5 min for 100 wavelengths and thousands of spectra with a global illumination instrument [11]. The speed of a line-scan spectrometer is somewhere between point mapping and global illumination depending on the number of pixels in the detector, but the number of wavelengths that can be acquired is fixed by the size of the detector.

On the other hand, due to its smaller area, a single pixel element of a 2D detector integrates less signal than a single-element detector used in a mapping study. Therefore, the signal to noise ratio (SNR) is lower in global imaging than in mapping when the same detector illumination is used. For example, in FTIR global imaging, the noise due to the FPA exceeds all other sources of noise (microscope or spectrometer), and integration time and the number of frames coadded must be optimized to achieve a good SNR in a relatively fast acquisition time [35]. Alternatively, a method of reducing noise for MIR imaging systems has been proposed based on the acquisition of several backgrounds [36]. The same holds for global Raman imaging: the laser has to be defocused to cover a large surface area, reducing the power reaching each element of the detector. Spatial resolution, acquisition time and spectral resolution have been compared between point mapping, line mapping and global imaging [37,38]. One author concluded: "With constant power density, the global imager is much faster than a line or focus point when an image with only one Raman shift is acquired but much slower if a complete spectrum is acquired at a single sample position" [37,38].

During a global imaging experiment, other sources of noise due to the optics and the nonuniform pixel illumination must also be taken into account. Different calibration schemes may help to decrease noise due to nonuniform illumination (Section 2.5).

Point mapping gives spectral resolution twice as accurate as global imaging experiment [39] and is the best technique for studying the distribution of minor compounds [32]. Moreover, it is more appropriate for analyzing rough-surfaced samples because the focus can be tuned to each spatial position. The consensus is that the choice of technique depends on the application. Point mapping is ideal for analyzing a small area for minor compounds (e.g. impurity detection). If the SNR is less critical, and an overview of a well-characterized sample is required, then global imaging is more appropriate. Choosing whichever technique is a matter of compromise.

### 2.4. Sampling techniques

#### 2.4.1. MIR and NIR sampling techniques

*2.4.1.1. Transmission.* In transmission measurement, the source illuminates the sample and the detector is placed behind the sample (Fig. 4) to acquire the fraction of light transmitted through the sample. Transmission analysis requires the sample to be partly transparent. In most cases, in the MIR range, samples must be diluted in nonabsorbing matrix otherwise no light might be transmitted to the detector. Liquid can thus be prepared as a dilute solution in a cell. Solid samples are dispersed in a potassium bromide disk or mull. Moreover, the powder particle size must be smaller than the radiation wavelength to avoid the Christiansen scattering effect which appears as band distortion in the spectra [14]. Transmission has been extensively used to analyze thin samples such as films [40–42] or tissues [43–45]. It is not possible with thick samples such as tablets.

*2.4.1.2. Reflection.* In reflection measurement, the detector is placed on the same side of the sample as the source to record the signal reflected by the sample. The sample is presumed infinitely thick and incapable of transmission. The two types of reflection measurement commonly used in CI analysis of pharmaceutical forms are diffuse reflection (DR) and ATR.

*2.4.1.2.1. Diffuse reflection.* Incoming radiation interacts with the sample and is scattered by interaction with the particles. A fraction of this light is reflected by the sample and recorded by the detector (Fig. 5). In the MIR range, DR requires the sample to be diluted between 10 and 100 times to avoid saturation and band distortion [31]. For this reason MIR-DR is rarely used for imaging: its use has not been reported in the literature. On the other hand,

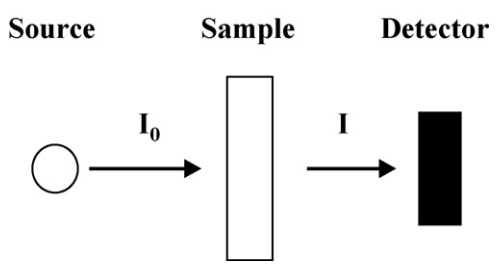


**Fig. 4.** Transmission measurement: the detector registers the signal passing through the sample.

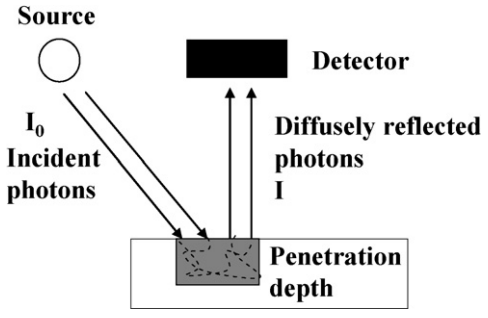


**Fig. 5.** Diffuse reflectance measurement and penetration depth.

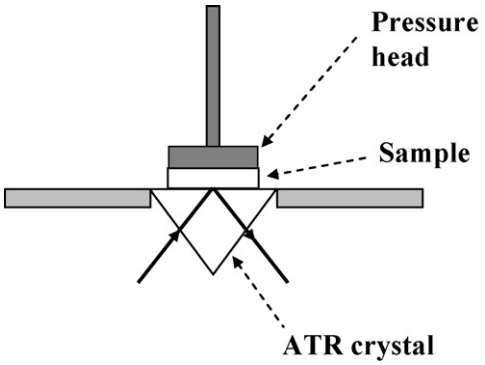


**Fig. 6.** Attenuated total reflection (ATR). The sample is placed on the surface of a crystal. A pressure head ensures good optical contact.

because samples need no dilution at all in the NIR range (the bands are weak), NIR-DR is widely used for the image analysis of thick nontransparent samples in various noninvasive applications [6,46], notably in the food industry [8,47,48] and pharmaceutics [49,50], e.g. tablets.

*2.4.1.2.2. Attenuated total reflection.* ATR was developed to avoid sample dilution in MIR analysis [14]. When a beam of light passes from a medium of high refractive index $n_2$ to one of lower refractive index $n_1$, with an angle $\theta$ greater than a critical angle $\theta_c$ ($\theta_c = \sin^{-1}(n_1/n_2)$), total internal reflection of the light occurs. ATR sampling relies on this property of light. The sample is placed on the optical surface of an ATR crystal objective (Fig. 6), usually made of zinc selenide (ZnSe), diamond, silicon (Si) or germanium (Ge) with refractive indices of approximately 2.4, 3.4, and 4.0. As the refractive index of organic material is around 1.5, the light is totally reflected at the crystal–sample surface interface. However, the electric field of the radiation, or evanescent wave, penetrates a short distance (a few micrometers) and interacts with the sample. The penetration depth is defined as the depth $E_0$ at which the electric field amplitude falls to a value $E = E_0 \exp(-1)$ and is given by the following equation [51]:

$$d_p = \frac{\lambda}{2\pi n_2 \sqrt{\sin^2\theta - (n_1/n_2)^2}} \quad (2)$$

The penetration depth is dependent on the wavelength λ. Thus only a minute sample volume is actually measured. This avoids scattering effects [51,52] and allows analysis of solid samples, such as powder or tablets, without need for dilution [53]. However, good optical contact must be ensured at the crystal–sample interface. Contact is readily achieved with soft materials, but when analyzing solid dosage forms, pressure needs to be applied, which may damage the sample or crystal. Diamond crystals can withstand high pressures but cost limits their size. ZnSe crystals have become a common alternative [31].

*2.4.1.3. Accessories for high-throughput analysis.* One advantage of an FPA detector is that several samples can be studied simultaneously in a controlled environment using dedicated sampling accessories.

Snively and Lauterbach proposed several sampling accessories for high throughput of gases in transmission analysis [54], e.g. a seven-sample temperature- and pressure-controlled reactor to test catalyst heterogeneity at high temperature, and a 16 stainless steel tube assembly sandwiched between MIR transparent windows for monitoring 16 gas-phase reaction products.

In several publications, Kazarian et al. have reported using polydimethylsiloxane (PDMS) grids placed on the surface of a macro-ATR crystal to analyze 156 liquids on a field of view (FOV) measuring approximately 1.5 cm × 2.2 cm [55]. Another four-channel PDMS grid is used to study four dissolutions simultaneously [56]. Other cells are designed to control temperature and humidity [40] for the investigation of powder compaction [57].

*2.4.2. Raman sampling techniques*

Raman analysis is much more straightforward than IR spectroscopy and little sample preparation is needed. Visible and NIR laser sources allow analysis through cell or glass, and with the confocal aperture the laser beam penetrates only few micrometers into the sample. Moreover, sample morphology does not affect band shapes or relative intensities [31]. However, the heat induced by focusing the laser beam at high resolution may burn the sample.

## 2.5. Calibration

As in classic spectroscopy, calibration is a mandatory preliminary to hyperspectral imaging. In particular, spatial variations due to nonuniform illumination and detector noise must be reduced when using an FPA.

*2.5.1. MIR and NIR calibration*

The parameter of interest in IR spectroscopy is the fraction of light absorbed by the sample. Absorbance is indirectly calculated using the intensity of incoming radiation (background signal $I_0$), and the light intensity recorded either in transmission or reflection. In transmission, acquisition of the background signal is straightforward, being recorded when there is no sample in the beam path. In reflection, a glass surface with a fine gold coat reflects the totality of the MIR beam. With ATR, the crystal is simply placed in air. Once a sample signal has been recorded in transmission or reflection (intensity $I$), sample absorbance $A$ is usually calculated using the following equation:

$$A = -\log\left(\frac{I}{I_0}\right) \quad (3)$$

In NIR-DR spectroscopy, background is typically acquired using a high reflectance ceramic standard. The absorbance spectrum is then calculated using Eq. (3). In NIR-CI, however, InGaAs detectors are often used. They have a dark current – the detector response when no signal is recorded – that is wavelength-dependent [58]. It is measured by blocking the camera lens and must be subtracted from the raw signal. Reflectance is then calculated using the following equation:

$$R = \frac{I - D}{I_0 - D} \quad (4)$$

where $D$ is the dark response of the detector.

Absorbance $A$ is given by

$$A = -\log\left(\frac{I - D}{I_0 - D}\right) \quad (5)$$

DR calibration with acquisition of a dark cube and 100% standard reflectance assumes detector response to be linear between 0% and 100% reflectance. However, this assumption may be mistaken and finding a true 100% reflectance standard may be difficult. More robust calibrations are described in the literature. Proposals have included the use of four standards with 2%, 50%, 75% and 99% reflectance [58], fitting a linear or quadratic regression at each pixel position using the four values (quadratic regression gave a better fit in that study, suggesting that the detector response was nonlinear). The same group have also used six standards (dark, 2%, 25%, 50%, 75%, and 99%) and three methods of calibration [59]: (1) global or pixelwise data subset selection; in global selection a

median spectrum is computed for each reflectance standard hypercube, regression is performed, and each pixel is corrected using the same coefficients; in pixelwise section, each pixel is individually corrected; (2) a linear or quadratic regression model, as previously proposed [58]; and (3) comparing the possibility of first scanning each standard and then the sample (external calibration) or integrating the standards in the same FOV with the test sample (internal calibration); external calibration corrects pixel-to-pixel variance whereas internal calibration may correct time-dependent variations due to temperature or power fluctuation. The first experiment tested the value of scanning several standards for calibration. The results clearly demonstrated the improvement in calibration quality with the combination of a range of spectral standards, fitting to a quadratic model, and pixelwise correction. The second experiment simulated lamp aging (decreasing lamp power); internal correction reduced measurement error.

Another standardization method has been recently proposed for line-scan NIR imaging systems [60]. Calibration is first performed following Eq. (4). Several reflectance images of samples giving a spatially uniform spectral response are then acquired. Since the sample spectral response is uniform, most pixel noise is presumed due to illumination unevenness or difference among detectors. To correct for these artifacts, data cubes ($x \times y \times \lambda$) are first reduced to an $x \times \lambda$ 2D image (termed "average line image" by Liu et al.) by averaging the $y$-dimension, thereby relating each value of the 2D image to a detector sensor. The average spectrum of each data cube is then calculated and used as a reference value. The spatial/spectral values of $l_{x\lambda}$ (average line image at the spatial/spectral coordinate position $x$ and $\lambda$) for all average line images are plotted on the same graph (24 average line images in this case) against their reference value, revealing a linear relationship from which slope and bias are calculated. Each spectral/spatial position of the detector is thus associated with two correction coefficients (slope and bias) which are then used to correct new data cubes. This method is easy because only homogeneous standards are required and the 'true' reflectance does not need to be known. The standardization performed in this study significantly decreased the standard deviation (S.D.) at each wavelength band.

#### 2.5.2. Raman calibration

As Raman spectroscopy involves light scattering, not absorbance, acquisition of the Raman spectrum is direct and does not need to be preceded by recording a background signal. The calibration of Raman systems is not required before each measurement. Silicon or emission lines from light-emitting diodes generate strong, sharp and well-resolved peaks at a particular Raman shift. Thus checking the peak positions of these standards confirms apparatus Raman shift accuracy from day to day. A standard with known intensity over the wavelength (such as a tungsten source or luminescent standard) might be used to correct for peak intensity [61]. Raman calibration can degrade over time due to vibration or temperature changes, but specifications determining when the device should be calibrated are rarely available [23].

### 2.6. Spatial resolution

In imaging it is important to know the size of the finer patterns to be resolved by the apparatus. This is determined by its spatial resolution. Spatial resolution is defined by the Rayleigh criterion as the minimum distance $d_m$ between distinguishable objects in an image.

Spatial resolution in hyperspectral imaging may be limited by two factors: on the one hand, device optics and the diffraction limit imposed by the wavelength, and on the other, the depth of radiation penetration into the sample, itself dependent on the measurement mode.

#### 2.6.1. Diffraction limits

For a microscope, $d_m$ can be defined by the Rayleigh criterion in the following equation:

$$d_m = \frac{1.22\lambda}{2n\ \sin(\theta)} \quad (6)$$

where $\lambda$ (nm) is the wavelength, $n$ the refractive index of the medium, and $\theta$ the collecting angle of the device. $n \times \sin(\theta)$ is also known as the numerical aperture (NA).

There have been several studies of spatial resolution in MIR. One study employed a standard USAF 1951 resolution target to compare the spatial resolution achievable in transmittance by several imaging spectrometers [43]. The best imaging system gave a spatial resolution down to 2.4 μm at 4000 $cm^{-1}$, according to the Rayleigh criterion using a 36× Cassegrain objective. Other studies have assessed the spatial resolution achievable in reflection using micro- and macro-ATR accessories and a nylon filter or polymer film with fine patterns [33,62–64]. Micro-ATR may be able to achieve 4 μm spatial resolution in reflection. Macro-ATR may be useful for analyzing large sample areas and gives a spatial resolution of ~10 μm. However, macro-imaging generates spatial distortion of the image which must be corrected.

In Raman microspectroscopy, spatial resolution is limited by the size of the laser beam in mapping experiments but can attain 1 μm. In-line mapping and widefield Raman imaging, submicron resolution is achievable (e.g. 250 nm using 514.5 nm laser excitation for global imaging) [34,65].

In order to overcome physical and instrumental limitations and achieve yet higher spatial resolution, some authors have developed techniques known as super-resolution [66]: several low-resolution images of the sample are first acquired, then precisely known translational shifts between these images produced by a computer-controlled stage are mathematically analyzed to produce the high-resolution image. Such techniques open up new opportunities in submicron pattern CI.

#### 2.6.2. Penetration depth

According to the photon pathlength in DR measurement mode (Fig. 5), a photon entering the sample will interact with the particles and several reflections and diffractions will appear before it exits the sample and reaches the detector. The gray area indicates the volume in which interaction may occur. Penetration depth is a measure of light entry into the sample. The shorter the wavelength, the higher the energy of the incident beam. High energy leads to deeper penetration and thus to an increased volume of interaction. Penetration depth is thus the limiting factor in NIR-DR imaging [67]. A study using varying thicknesses of cellulose placed on top of a substrate found that shorter wavelengths (1100 nm) penetrated to an information depth of as much as 777 μm whereas longer wavelengths (2380 nm) penetrated to 109 μm [68]. The maximum resolution achievable with NIR-DR is probably not less than 30 μm due to the penetration depth of NIR radiation [67].

## 3. Analysis of hyperspectral imaging data

A key issue in hyperspectral imaging is how to extract relevant information from the huge amount of data. Visual inspection of each wavelength image is time-consuming and probably impossible when attempting to compare several data cubes. The data must therefore be reduced by several image planes and quantitative

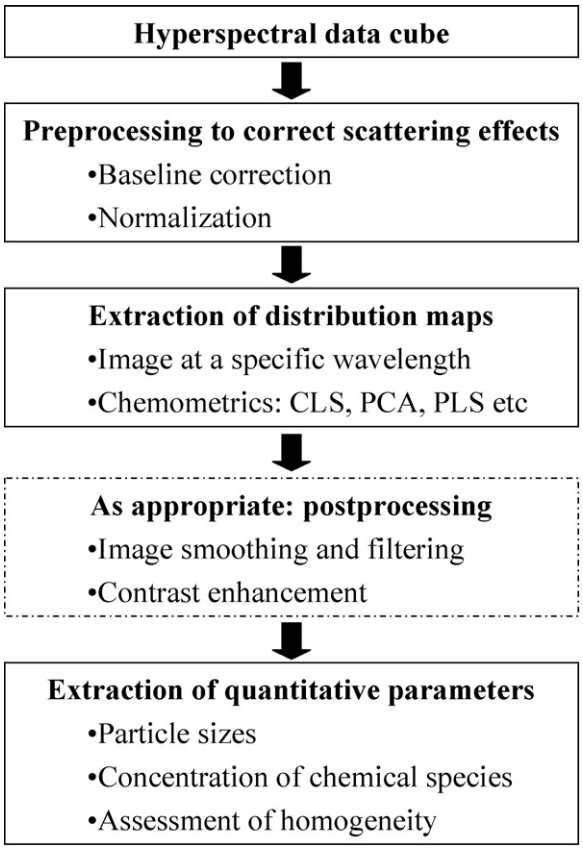


**Fig. 7.** The four steps in the processing workflow of a hyperspectral data cube, with examples of methods. CLS: classical least squares; PCA: principal component analysis; PLS: partial least squares.

parameters. In the classical processing workflow of an hyperspectral data cube (Fig. 7), the first step is preprocessing, to reduce baseline and scattering effects. The next step is to spatially locate each compound. Post-processing can also be applied to enhance contrast or reduce noise. The final task is to determine parameters, such as particle size and shape and the distribution of pixel intensity, which enable the images to be objectively compared. As hyperspectral imaging is an application spanning spectroscopy and image processing, most processing methods derive directly from these fields and have already been described in numerous reviews [1] and textbooks [69–71]. Their detailed exposition thus falls outside the remit of this review, which will focus instead on their CI applications.

### 3.1. Preprocessing

Spectral and spatial artifacts, such as rough surfaces, optic effects, and detector noise, must be removed in the preprocessing of a hyperspectral data cube. Hyperspectral raw images mostly feature the same patterns depending on wavelength and reveal physical differences on the tablet surface rather than chemical variation. Appropriate preprocessing enables compound distribution maps to be visualized [39]. The spectral pretreatments used in classical single-point spectroscopy can be applied to hyperspectral imaging. The first preprocessing step is baseline correction to reduce the background effect of uncontrolled spectral variation. Derivative preprocessing is one solution but it increases the noise. Other techniques use polynomial curves [72]. The second important step in preprocessing is normalization. Raw spectra often feature overall intensity variation due to optical path differences [73]. Normalizing the spectra consists of dividing each spectral value $x_\lambda$ of a spectrum $x$ by a representative number $a$.

$$\text{xnorm}_\lambda = \frac{x_\lambda}{a}, \quad \text{for } \lambda = 1 \ldots \lambda_{\max} \tag{7}$$

where $a$ can be the maximum value of the spectra, or the sum of all values in the spectra (unit area) or the sum of squares of all values (unit length). In the event of unknown variation (offset) between spectra, it can be useful to subtract the mean value $\bar{x}$ of $x$ spectra from each spectral variable $x_\lambda$.

NIR spectra are often normalized using the standard normal variate (SNV) [74]. The denominator of the equation is the S.D. of spectrum $x$.

$$\text{xnorm}_\lambda = \frac{x_\lambda - \bar{x}}{\sqrt{\sum_{k=1}^{p}(x_k - \bar{x})^2/(p-1)}}, \quad \text{where } \bar{x} = \sum_{\lambda=1}^{\lambda_{\max}} x_\lambda \tag{8}$$

A final correction to remove scattering effects can also be applied. Multiplicative scatter correction (MSC) [75] was developed for that purpose. It linearizes the data to a certain extent. An ideal spectrum representative of the data set $X$ is first chosen. The mean spectrum $\bar{x}$ is a good approximation. Linear regression is then performed between each spectrum $x_i$ of the data set and the representative spectrum, leading to the calculation of the slope ($b_i$) and bias ($a_i$).

$$x_{i\lambda} = a_i + b_i\overline{x_\lambda} + e_{i,\lambda} \tag{9}$$

The corrected data are then given by

$$\text{xcor}_{i\lambda} = \frac{x_{i,\lambda} - a_i}{b_i} \tag{10}$$

Classical spectral preprocessing is thus applied to hyperspectral data. However, imaging has an advantage over classical spectroscopy: comparing the remaining pixel-to-pixel spectral variations can give more information about the sample or about the ability of the preprocessing method to clean the spectra. For example, in a study to determine which spectral method was most effective at reducing baseline and scattering effects due to different sizes of salt and sugar particles [76], each sample was scanned individually. Kubelka Munk, SNV and absorbance transforms, unit length and unit area normalization, first and second derivative, and several MSC variants [75] were then applied to reduce the scatter effects in the hyperspectral NIR images. MSC was first applied to individual images, then to all images simultaneously, thus leading to different representative spectra for use in correction. Principal component analysis (PCA) (Section 3.2.2.1) showed that global piecewise MSC removed all variation due to particle size within spectra, whereas derivative spectra did not remove particle size dependency. Next, a partial least squares (PLS) model (Section 3.2.2.1) was computed to predict sugar particle sizes. PLS predictions based on the mean spectra provided accurate results, but predictions made on individual pixels located differences in particle size within a single tablet. Although global piecewise MSC is best at minimizing particle size effects, PLS was still able to find sufficient correlated variance for accurately determining particle size, as it led to broader particle size distribution. Predictions based on mean spectra did not yield such information.

Kohler et al. [77] applied extended multiplicative signal correction (EMSC) [78] to separate physical and chemical information in FTIR microscopy images of biological tissues. The spatial information provided by hyperspectral imaging showed that the scatter changes were in fact induced by textural changes resulting from the heat treatment of beef, and that spectral pretreatment was thus essential for correctly interpreting the spectral changes due to the chemical nature of the sample.

More advanced methods such as wavelet preprocessing can also be applied to hyperspectral imaging to reduce noise in Raman imaging spectra [79,80]. In these studies, prior spectral denoising using

wavelet analysis on the spectral range increased the correct pixel classification rate. Wavelets not only removed noise but also compressed the data. They may thus help to reduce the memory load and computing time needed for processing hyperspectral images.

Pretreatments can be applied to the spectral as well as spatial dimensions. Including both dimensions may be more effective in reducing noise and revealing fine spatial patterns.

Transformations available for removing noise in the spectral and spatial dimensions as applied to FTIR hyperspectral imaging [81] begin with baseline correction and band ratio methods in the spectral dimension. Transforms including PCA (Section 3.2.2.1), a modified version of the maximum noise fraction (MNF) [82], and the 2D fast Fourier transform (FFT) are then applied to the entire data cube. These techniques project the data into a new space using mathematical constraints. In this new space some elements are kept and transformed back into the initial space as partially denoised data. PCA and MNF optimize a criterion (maximum variance for PCA and between-band correlation for MNF) which leads to a set of eigenvalues. The largest eigenvalues are considered as describing relevant information whereas the smallest ones are noise-related and discarded in the reverse transformation. FFT is used to transform each data cube image in the Fourier space where higher-frequency components that are presumed to be noise are rejected from the center. A spatial mask then cuts out these components. Central values are projected back into the image space as a smoothed image. In these studies, MNF and PCA discriminated between noise and relevant information whereas FFT led to the loss of fine spatial patterns of interest.

In another example of combined spectral/spatial correction [83], the FFT was applied first to the spectral dimension and then to the spatial dimension to minimize pixel-to-pixel baseline interference in FTIR imaging data. This successfully reduced baseline variation and data noise compared to classic two-point baseline flattening correction.

Spatial information is also useful in preprocessing Raman spectra in order to remove random high-intensity spikes arising from cosmic rays. Classic Raman spectroscopy achieves this by summing several spectra of the same sample [84]. However, because a single imaging experiment acquires many spectra, alternative methods have been proposed for this special case. For example, a procedure based on correlating neighboring pixels in an image data set assumes that a pixel has at least one adjacent pixel having similar spectral information; thus if a pixel does not have a high correlation with at least one of its adjacent pixels, it can be assumed to be a spike [85]. Zhang and Henson [86] subsequently considered the case of samples with a low API content, where one pixel may hold an API signal but its neighbor not. Whereas the previous method might consider this pixel an outlier, their proposed alternative is to add the reference spectra of the pure compounds into the regression model in order to avoid pixel misclassification.

Wavelet preprocessing can also be applied to both dimensions. Vogt et al. compared several techniques for hyperspectral imaging, showing that Haar wavelets give the fastest computation timed, highest compressions, and preserve more spatial and spectral information, whereas Daubechies wavelets preserve spectral information [87,88]. They also proposed a hybrid algorithm combining different wavelets for spatial and spectral compression; the best combination in terms of time saving and information preservation is Daub8 for the spatial dimension and Daub4 for the spectral dimension [88].

### 3.2. Extraction of distribution maps

Chemical compounds in the sample are localized by extracting distribution maps. Extraction has to be as accurate as possible to avoid pixel misclassification. Many methods have been developed (Fig. 8), most of which derive directly from classic spectroscopy or image processing. Choice of the most appropriate method is based on the information available, the spectral signatures of the pure compounds in the study system, and experimental noise. The major division within these methods is that between univariate and multivariate analysis.

#### 3.2.1. Univariate analysis

Univariate analysis is the simplest method of obtaining distribution maps. Laser light is absorbed (MIR and NIR) or scattered (Raman) by each chemical entity in a mixture at specific wavelengths depending on its chemical bonds. Selecting the image at those wavelength positions localizes the specific compounds depicted by the pixels with the highest absorbance, i.e. highest intensity.

An alternative method is peak integration. At each pixel position, the area under a specific spectral peak is calculated. The resulting values form a new image in which higher pixel intensity reflects higher compound concentration.

If these methods are the most straightforward in revealing chemical localization, they also require the study system to be well characterized. All compounds should be known and raw powders must be scanned to characterize their spectral signature. Since the spectroscopist often knows the composition of the medicine, univariate analysis may suffice in most studies. For example, univariate analysis has been used to process FTIR imaging data sets [40–42,56,57,62,89–92]. However, overlaps in complex systems may prevent identification of specific wavelengths for each compound, especially in the NIR range. In addition, there will be cases where a constituent is unknown (contamination studies, for example) or at such low concentration that univariate analysis fails to extract sufficiently reliable distribution maps. In such cases, multivariate analysis is more appropriate for characterizing fine spectral variation and for localizing compounds [32,93,94].

#### 3.2.2. Multivariate analysis

Multivariate analysis takes into account all the spectral information contained in the data cube. Many analytical pathways are available (Fig. 8), from factor analysis methods that aim to decrease dimension by using an underlying multivariate distribution [95,96], modeling the data into a linear combination of factors plus a noise term, to clustering techniques seeking to group spectra (whether full or in the reduced spatial dimension after factor analysis) on the basis of shared features [97].

*3.2.2.1. Factor analysis.* Factor analysis of hyperspectral data is based mainly on algorithms derived from classic spectroscopy in which each row of a 2D matrix represents the spectra of one measure. However, instruments generate a stack of images, i.e. a 3D matrix. An unfolding step is thus required before classic factorial methods can be applied. The data cube **D** of dimension $n_{max} \times m_{max} \times \lambda_{max}$ is thus transformed into a 2D matrix $\mathbf{X} = ((n_{max} \times m_{max}) \times \lambda_{max}) = i_{max} \times \lambda_{max}$ (Fig. 9). The processing of these unfolded cubes will be discussed first, before considering the methods analyzing three-way or higher dimensionality matrices that have been developed in recent decades by the chemometric community.

*3.2.2.1.1. Two-way processing. Principal component analysis and partial least squares.* PCA and PLS [95] are the two algorithms most used in the chemometric community for extracting chemical information. PCA aims to reduce matrix dimensionality by removing correlation between variables. The data are projected on a new principal component subspace (Fig. 10). The principal components are iteratively computed, the first being constructed to explain as

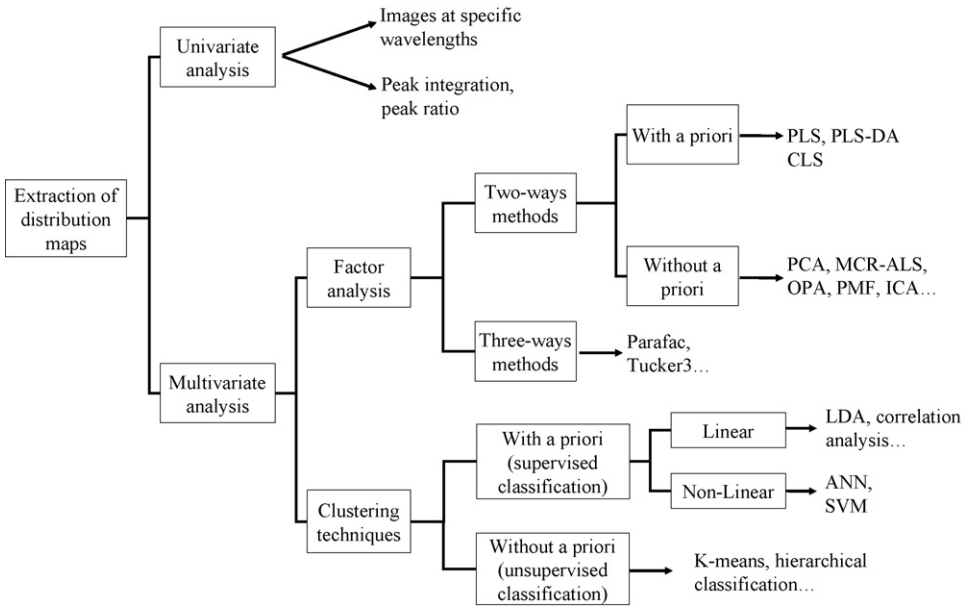


**Fig. 8.** Classification of methods for extracting distribution maps and some examples. ANN: artificial neural network; CLS: classical least squares; LDA: linear discriminant analysis; MCR-ALS: multivariate curve resolution-alternating least squares; OPA: orthogonal projection analysis; PCA: principal component analysis; PLS: partial least squares; PMF: positive matrix factorization; SVM: support vector machine.

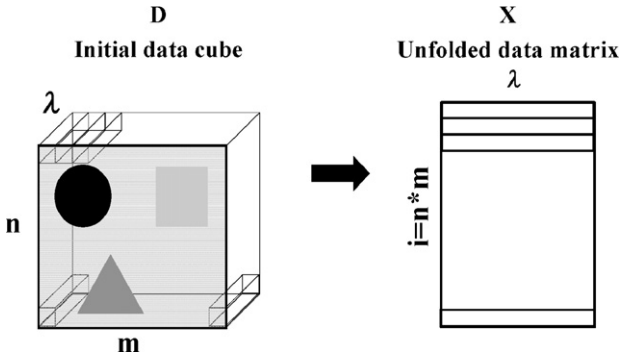


**Fig. 9.** Unfold procedure. The three-dimensional matrix is transformed into a two-dimensional matrix by unfolding the spatial dimensions.

much of the data variance as possible, the second being constrained to be orthogonal to the first and explaining the residual variance not taken into account by the first, and so on. The principal components are known as loadings and data projections onto the principal components are known as scores. The first loadings account for the most useful information whereas later loadings mostly describe noise. Dimensionality is reduced by discarding these noisier components. After the PCA transform, each column of the score matrix is folded back to form an image that represents pixel variability along the corresponding loadings. PCA has been extensively used in hyperspectral image analysis to determine compound distribution in NIR-CI [50] and in particular to identify minor compounds in Raman imaging [32,65,98], since it can reveal fine spectral variations. However, while useful in explaining variance, it does not have chemical meaning. Since PCA loadings depict negative values, it may be difficult to relate them to chemical compounds.

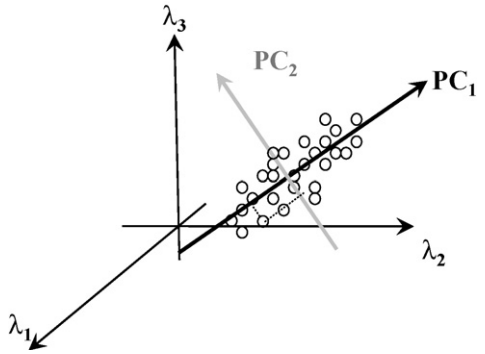


**Fig. 10.** Geometric visualization of principal component (PC) analysis. The first principal component captures most of the variance. The second is constrained to be orthogonal to the first and to explain the residual variance not taken into account by the first.

PLS [99,100] is used to predict the variable(s) of matrix **Y** based on observing the variable(s) of matrix **X**. **Y** is projected onto latent variables $(t_1, t_2, \ldots, t_{k_{max}})$ which are linear combinations of $x_1, x_2, \ldots, x_{\lambda_{max}}$. A first calibration step is required to build the mathematical model linking the **X** and **Y** matrices, using a set of samples in which the **X** and **Y** matrices are clearly known. For example, in the case of NIR quantification of chemical compounds, the spectra form the rows of the **X** matrix and the known concentrations the rows of the **Y** matrix. Once developed, the model must be validated with a set of known values before it can predict unknown samples. Burger and Geladi [101] recently compared spectra from a hyperspectral image and those from two classic spectrometers in terms of sample content prediction. The prediction error from the hyperspectral imaging spectra was between that of the two spectrometers, showing that the hyperspectral image can be used for quantification. PLS regression applied to hyperspectral imaging has at least two advantages over classic spectroscopy: (1) after predicting each pixel, each column of the predicted **Y** matrix can be folded back to display the spatial distribution of predictions; for example, when the predictions are relative concentrations, the compound distribution maps can be visualized [101,102]; (2) a data cube might contain several thousand spectra which can be divided between calibration and validation sets; the calibration models can be computed using the mean (or median) spectrum of individual images. The user might also consider the mean spectrum of several spatial regions of interest (ROI) in a data cube including thousands of spectra as a calibration set [102]. The data can then provide

several calibration sets to optimize the model, for use in predicting pixel-to-pixel concentrations.

In hyperspectral imaging, it is also possible to construct a PLS-discriminant analysis (PLS-DA) classification method using pure reference spectra to extract the compound distribution maps. The aim is to identify latent variables that will enable class separation by taking into account the class membership of observations [103]. This method has been used to predict binary mixtures from NIR-CI images [104] but also to display compound distribution [11,39,105]. However, when a complex matrix is under study or when compounds are homogeneously distributed or present at low concentration, extracting distribution maps by PLS-DA classification may be less straightforward, due to differences between the training library (the pure spectra) and the sample to be predicted [39].

*Bilinear modeling*. The two classic factorial methods, PCA and PLS, use mathematical constraints to extract the factors carrying the most relevant information. Although useful in explaining variance, however, it can be difficult to interpret the loadings and relate them to a chemical species. Thus alternative methods have been developed to unravel the pure spectra of the chemicals and their respective concentrations from the mixed spectra. The analysis is based on the bilinear model, according to which a sample's absorbance results from the sum of the absorbances of its component chemical species. Thus the mixed spectrum may be viewed as the weighted sum of each pure material spectrum plus experimental noise. This phenomenon is mathematically described in the following bilinear model:

$$\mathbf{X} = \mathbf{C}\mathbf{S}^{\mathrm{T}} + \mathbf{E} \tag{11}$$

where $\mathbf{X}$ is a two-way observed signal matrix, $\mathbf{C}$ is a column-wise matrix ($n_{\max} \times k_{\max}$) of chemical species concentration/abundance, $\mathbf{S}^{\mathrm{T}}$ is the pure spectra matrix ($k_{\max} \times \lambda_{\max}$), and $\mathbf{E}$ is residual noise (Fig. 11).

If the pure spectra matrix $\mathbf{S}^{\mathrm{T}}$ is available, the direct classical least squares (DCLS) algorithm, also known as ordinary least squares (OLS), may recover the distribution maps. This method consists in minimizing sum of the square errors, i.e. minimizing $||\mathbf{X} - \mathbf{C}\mathbf{S}^{\mathrm{T}}||^2$. $\mathbf{C}$ is estimated by the pseudoinverse:

$$\mathbf{C} = \mathbf{X}\mathbf{S}(\mathbf{S}^{\mathrm{T}}\mathbf{S})^{-1} \tag{12}$$

If a set of concentrations is available, CLS can also be used to estimate the pure spectra:

$$\mathbf{S}^{\mathrm{T}} = (\mathbf{C}^{\mathrm{T}}\mathbf{C})^{-1}\mathbf{C}^{\mathrm{T}}\mathbf{X} \tag{13}$$

To a certain extent DCLS may give information about relative constituent concentrations [39]. Although differences in spectral background and nonlinearity, especially in the NIR region, may prevent accurate concentration determination, DCLS extract reliable distribution maps for well-characterized samples, such as those from pharmaceutical development [34,93].

If sparse information is available for the analysis, then the $\mathbf{C}$ and $\mathbf{S}^{\mathrm{T}}$ matrices must be extracted simultaneously. Factorization of the $\mathbf{X}$ matrix into two positive matrices, $\mathbf{C}$ and $\mathbf{S}^{\mathrm{T}}$, can have an infinity of solutions. Mathematical constraints or constraints based on physical phenomena must therefore be introduced to reduce the space of possible solutions [106,107]. The analytical community refers to the methods that unravel mixed spectra as self-modeling curve resolution (SMCR). Many methods have been proposed since the 1970s [108] for factorizing the $\mathbf{X}$ matrix (reviewed in Refs. [109,110]). The first kind of SMCR algorithm factorizes the $\mathbf{X}$ matrix using mathematical constraints.

*Mathematical resolution*. To achieve loadings more representative of chemical information, methods have been developed that start in the principal component space and attempt to find appropriate rotation of the principal components. For example, key-set factor analysis (KSFA) [111] constructs the set of key factors that are the most orthogonal to each other, while iterative target transformation factor analysis (ITTFA) [112] tests whether the factor produced by principal component rotation is a true factor.

The recently proposed band target entropy minimization (BTEM) algorithm [113] has been applied to Raman images of pharmaceutical tablets [114]. Entropy is a measure of system disorder. If the entropy value is low, the system is organized and simple. Since pure spectra are assumed to be the simplest underlying patterns, minimization of entropy should be an appropriate choice. In the proposed procedure, each pure spectrum is individually estimated by rotating the principal components. After PCA, the rotation matrix is optimized by minimizing an objective function that includes two terms: the first minimizes entropy, while the second ensures non-negativity of the spectra.

Other methods work directly with the $\mathbf{X}$ matrix. Perhaps the most popular are the simple-to-use interactive self-modeling mixture analysis (SIMPLISMA) algorithm proposed by Windig and Guilment [115] and orthogonal projection analysis (OPA) [116]. SIMPLISMA identifies the purest columns in the $\mathbf{X}$ matrix using a criterion calculated from the column means and S.D.s. The purest columns form the $\mathbf{C}$ matrix, while the $\mathbf{S}^{\mathrm{T}}$ matrix is calculated using Eq. (13). OPA identifies the spectra showing most dissimilarity in the $\mathbf{X}$ matrix.

Independent component analysis (ICA) [117] has become particularly popular in recent years. It optimizes a criterion that assumes the sources to be statistically mutually independent. However, the basic assumption of independence may not always be fulfilled in this case because pure spectra (especially NIR spectra) may exhibit correlation [118,119]. Moreover, correlation between adjacent pixels may also prevent good separation by ICA.

*Introducing constraints based on physical meaning*. The above procedures produce hypotheses about the data or pure spectra in order to derive a mathematical resolution of the bilinear model. The solution is unique in most cases but it may happen that not all the spectroscopic data confirm the hypotheses. The procedures may also lead to unrealistic solutions, such as a negative spectral profile or concentration. Methods have therefore been developed that force the solutions into consistency with physical meaning. At the core of such methods is the definition of a criterion to be minimized under specific constraints. The constraints reduce the space for feasible solutions and can differ depending on the spectroscopic technique considered. The most widely implemented, because applicable to all kinds of spectroscopy, is that both the concentration and spectral profiles be positive. Non-negativity is the only constraint that can be set for hyperspectral imaging, with the possible addition of the number of compounds present at an individual pixel position, if known.

However, the currently most popular resolution method in analytical chemistry, including hyperspectral imaging, is undoubtedly multivariate curve resolution-alternating least squares (MCR-ALS) analysis [107]. The $\mathbf{X}$ matrix is factorized by minimizing the sum-of-squares norm of the matrix of error $\mathbf{E}$. The objective function $Q_1$ to be minimized is then given by the following equation:

$$Q_1 = \sum_{i=1}^{n}\sum_{j=1}^{m}\left\| x_{ij} - \sum_{k=1}^{p} c_{ik}s_{kj}^{\mathrm{T}} \right\|^2 \tag{14}$$

$Q_1$ minimization is achieved by alternately calculating the $\mathbf{C}$ and $\mathbf{S}^{\mathrm{T}}$ matrices by least squares fit (Eq. (12) and Eq. (13), respectively) while the other matrix is fixed. The constraints are applied with every iteration after estimating one product matrix. For example,

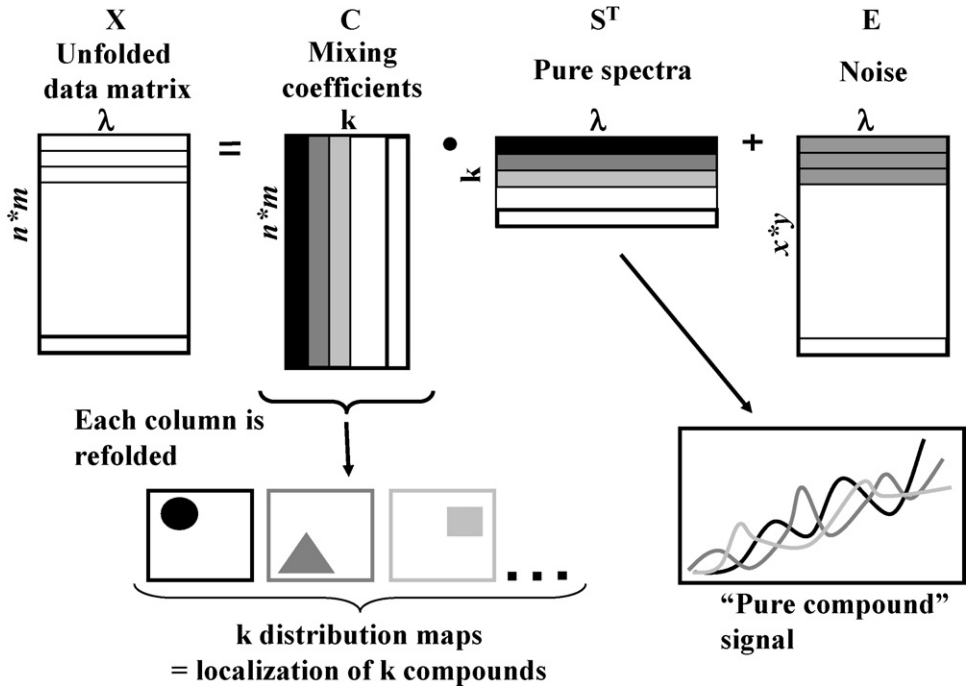


**Fig. 11.** Bilinear modeling: the mixed spectra matrix is factorized into two matrices related to concentration and pure spectra.

the positivity constraint can be forced by setting negative values to zero or applying the fast non-negative least squares (FNNLS) algorithm [120] for smoother constraint.

In 1999, in a letter to *Nature*, Lee and Seung [121] published the non-negative matrix factorization (NMF) for separating sources. The algorithm also minimizes $Q_1$ with specific update rules for the **C** and $\mathbf{S}^T$ matrices that ensure spectra positivity and provide a mixing coefficient for positive matrix factorization. A modified version of the NMF, constrained NMF (cNMF), that includes constraint on the minimum amplitude of the recovered spectra, has been applied to Raman imaging spectra [122] with good extraction results.

An alternative algorithm is positive matrix factorization (PMF) developed by Paatero [123] which minimizes the $Q_2$ criterion.

$$Q_2 = \sum_{i,j} \frac{(\mathbf{X} - \mathbf{C}\mathbf{S}^T)^2_{i,j}}{\sigma^2_{i,j}} \quad (15)$$

where $\sigma_{i,j}$ is the estimated uncertainty of the $(i,j)$th variable. The effect of introducing such weighting is that variables known with the greatest precision will have the greatest influence on minimizing the $Q_2$ function. Positivity constraints are applied using penalty functions. The multilinear engine (ME) software developed by Paatero can be used to optimize $Q_2$ [124].

For computing MCR-ALS, PMF and NMF, matrices **C** and $\mathbf{S}^T$ can be initialized by random values or by preliminary estimates calculated using mathematical solutions such as OPA, SIMPLISMA, or PCA. Methods based on chemical constraints can thus be viewed as a refinement of a preliminary solution. The main drawback of these methods is that different initializations might produce different results.

Several studies have compared the performance of these algorithms in terms of hyperspectral imaging demixing. Thus Duponchel et al. evaluated OPA, SIMPLISMA, PMF and MCR-ALS in hyperspectral image deconvolution, as well as several initializations of MCR-ALS [125]. A synthetic MIR hyperspectral data cube containing six polymers was generated, together with different levels of noise and spectral shift to test the robustness of the methods. Best results were obtained with OPA, followed by MCR-ALS, which was reported to be less sensitive to the signal-to-noise ratio. Andrew and Hancewicz [126] also tried OPA/ALS and PCA/ALS on Raman images. The two algorithms had the same ability to extract pure spectra and distribution maps, but OPA/ALS was faster. The present authors compared MCR-ALS, PMF, and NMF for demixing NIR imaging data [127]. PMF proved best at recovering pure spectra, but finding appropriate values for $\sigma_{i,j}$ may not be straightforward.

*Spatial constraints*. Introducing spatial constraints may increase the accuracy of pixel classification in homogeneous regions, especially if noise is present. With CI data sets, spatial information can be used to determine the number of species present in a pixel, as proposed in window evolving factor analysis (WEFA) [128], which selects adjacent pixels included in a window centered on a pixel. PCA on the window then determines the number of species present. This value is assigned to the center pixel, and the window is then centered on the next pixel, and so on. Iteration of the procedure on all pixels results in what is known as a rank image. Determining the number of chemical species in a pixel makes it possible to refine factorization by incorporating zeros in the concentration matrix provided that more information such as reference spectra is known for clear identification of the present or absent chemical species. De Juan et al. applied this method to Raman imaging data sets [129,130].

To improve extraction, a modified version of ALS introduces a matrix of weights that can be set in different ways: constant values, linear change with iteration number, and percentage of **S** and **C** at each iteration [131]. Alternatively, the weight matrix can be constructed using probabilistic class partition (Bayesian discriminant clustering) [132]. Thus the probability of compound presence is determined during the computation at each pixel and acts as a spatial constraint.

*3.2.2.1.2. Three-way modeling for hyperspectral imaging.* Three-way modeling exists since the 1980s and may be an alternative solution for describing both spectral and spatial information. Using multiway notation, a hyperspectral data cube may be viewed as an OOV matrix where O is an object mode (i.e. observation or, in

our case, spatial dimension) and V a variable mode (i.e. spectral information). Gurden et al. applied the multiway analysis algorithm to multivariate images (successive video images acquired over time) [133,134]. The study system was relatively simple and the parallel factor (PARAFAC) algorithm [135] was able to explain both features of the experiment. However, when complex spatial patterns are present, low-dimensional linear decomposition of spatial dimensions is not straightforward. A drawback with the PARAFAC algorithm is that the same number of components is used for both the O and V directions. In the only other study in the literature applying the multiway analysis algorithm to multivariate images [134], the Tucker3 algorithm [136] was used for hyperspectral imaging decomposition. It explained the spatial dimension with a higher number of underlying components. However, in OOV mode, multiway algorithms are mainly useful for compressing and denoising the data. The number of images before and after transformation remains the same. The study therefore concluded that unfolded methods (i.e. two-way models) are better suited to explorative data analysis and classifying OOV data, e.g. hyperspectral data, whereas multiway methods are more appropriate for analyzing OVV data.

*3.2.2.2. Clustering techniques.* Clustering techniques group pixels that share similar features according to a criterion. A group of such pixels is referred to as a class or cluster. Such classes or clusters can be created using a criterion based on distance, probability, and *a priori* information.

Clustering techniques come in two families: supervised and unsupervised. Supervised clustering needs a set of samples already assigned to a class. This serves as a training set to compute a model that will then be used to classify unknown samples. Unsupervised clustering classifies pixels without such a reference. Comprehensive reviews are available in textbooks [97,137] and journals [138,139].

*3.2.2.2.1. Supervised clustering.* The two widely used supervised clustering techniques in CI are linear discriminant analysis (LDA) and soft independent modeling of class analogies (SIMCA). LDA seeks a space that will best separate the class. The criterion to be maximized is given by the between/within group variance ratio. LDA has been mainly used to analyze IR maps of biological tissue [7,45]. In these studies, LDA was performed after selecting a spectral region using a linear discriminant guided genetic algorithm. Up to five kinds of biological tissue could be discriminated and the successful classification rate was 95%. SIMCA first computes a PCA for each class using the labeled spectra. For each PCA, limits are set defining the class in the principal component space. An unknown spectrum is then projected onto each class model and if the result falls between the limits of one of the classes, then the spectrum is labeled as belonging to that class [70]. Brain metastases have been accurately classified from IR spectroscopic images using SIMCA [44].

Other supervised classification methods are based on correlation analysis between pure spectra and pixel spectra. For example, cosine correlation analysis (CCA) uses the cosine of the angle between the pixel spectra and reference spectra in the $\lambda_{max}$ dimensional space. CCA of Raman spectra has been used to generate accurate tablet distribution maps [140].

Bayesian approaches may also be used for classification. Bayesian theory uses the probability that an observation may belong to class $\omega_i$. The model is constructed so as to minimize the probability of false classification. For example, Bhargava et al. [141] applied Bayesian classification to discriminate between 10 classes of cells and tissues in prostate histopathology. Instead of classifying the pixels by their spectral values, the authors first defined spectral metrics such as peak position, absorbance, and band ratio, then determined the number of metrics to be used to achieve the least classification error. Each observation was reduced to 20 metrics and yielded faster classification. The advantage of the Bayesian approach is that the calculation of the decision rules is not explicit and thus requires less computational load.

Another important linear method for supervised classification is multivariate image analysis (MIA) [142]. In MIA, PCA or PLS analysis is first applied to the data sets. Then a 2D score plot is generated. Since pixels having the same spectral features cluster in score plots, manual selection using a cluster training set can discriminate between pixel behavior. When the score plot region is correctly identified and linked to an image feature, pixels of new images can be classified using the model. At least two studies have used MIA for real-time process monitoring [143,144].

In the last decade, nonlinear approaches such as artificial neural network (ANN) and support vector machine (SVM) analysis have found applications in the processing of hyperspectral imaging data because they take nonlinear behavior into account. van den Broek et al. used multilayer feed-forward ANN (MLF-ANN) to identify plastic material by NIR-CI [145], while Lasch et al. classified FTIR microscopic imaging pixels using multilayer perception-artificial neural network (MLP-ANN) analysis for histological characterization [146]. Fernandez Pierna et al. used SVMs to classify compound feeds by NIR-CI [47,147]; comparison of SVM, ANN and PLS showed that SVM was the most accurate in classifying their data. Similarly, comparison of SVM versus MIA classification in discriminating splits and knots in NIR-CI images of lumber showed SVM to be an accurate automatic sampling technique [148].

*3.2.2.2.2. Unsupervised clustering.* Unsupervised clustering (reviewed in Ref. [149]) groups pixels having similar characteristics with no need for references or with at most information about the number of clusters. Perhaps the most popular example is the K-means (KM) algorithm and its fuzzy-C means (FCM) variant. KM is a distance-based hard classifier which assigns a pixel to one class only, whereas FCM attributes a degree of class membership. Mansfield et al. applied FCM to the analysis of NIR multispectral images of blood and tissue [150,151]. FCM has also been used to segment Raman spectral imaging data [80]. Two FCM algorithm variants were developed to detect small clusters in the analysis of multispectral visible images [152]: the cluster size insensitive fuzzy-C mean (csiFCM) algorithm outperformed classic FCM. Lasch et al. [153] compared the ability of KM, FCM, and hierarchical clustering to segment micro FTIR tissue. The agglomerative hierarchical (AH) algorithm proved the best method for differentiating tissue structure.

Šašic et al. proposed a clustering method based on 2D correlation [154,155]. After the mean spectrum is removed from the data set, the 2D correlation matrix is calculated and used to identify the most characteristic spectra. Distribution maps are generated using the CLS algorithm or univariate analysis. The method can be applied to Raman images of pharmaceutical samples.

Spectral identity mapping (SIM), a modified version of CCA, extracts distribution maps in an unsupervised manner, and has successfully characterized polymer-coated paper substrates [156].

The possibility of using spatial information to extract more relevant segmentation has also been tested. Several clustering algorithms were evaluated in combination with spatial information, including initialization, optimization, and post-processing [157]; not only did the inclusion of spatial information improve the segmentation of images containing homogeneous regions, it also appeared to reduce the problem of overlapping clusters and noise. Lin et al. have recently proposed a novel approach using a spatial directed agglomeration clustering method to detect drug polymorphs by Raman imaging [158].

### 3.3. Extraction of quantitative parameters

Having reduced the data cube and extracted the distribution maps, the next task is to develop methods for interpreting the images: because pattern perception in images is user-dependent, the application of image-processing techniques to enhance an object of interest can extract user-independent information. Such techniques (reviewed in Ref. [159]) include color and contrast enhancement, segmentation into homogeneous areas, edge detection, and texture classification.

#### 3.3.1. Image enhancement

Typically, images are digitized in 256 steps from low-intensity to high-intensity pixels. A color is assigned to each of the 256 values according to a look-up table (LUT). For example, a gray-level LUT assigns black to low-intensity pixels and white to high-intensity pixels. Intermediate values are displayed in shades of gray. A jet LUT assigns blue to low-intensity pixels and red to high-intensity pixels. Intermediate colors of the visible spectrum are used to display intermediate values. Color assignment is normally linear. However, to enhance regions with higher or lower intensities, square roots or log transfer functions may be used instead. Other techniques such as histogram equalization may also provide better contrast. Thus mathematical filtering and contrast adjustment of the histogram were applied to a single-band image [81]. Mathematical filtering changes the value of each pixel depending on the neighboring pixel values, and the different filters smooth the image or enhance the edge. Bhargava et al. explain the advantages and drawbacks of each technique and suggest that depending on the application, the spectroscopist might find one or other method more appropriate to their image data sets.

Another useful tool for synthesizing results is red–green–blue (RGB) reconstruction. Each pixel has a value between 0 and 255 for red, green and blue, generating $255^3$ possible colors. By assigning a compound distribution map to each plane of the RGB image, the localization of three compounds can be displayed simultaneously.

#### 3.3.2. Histogram analysis

Histograms are an important tool in image analysis. They plot pixel number against intensity. The relative maximum of a histogram is termed a mode. A distribution map histogram gives information about sample homogeneity. Thus a histogram showing a symmetric distribution with a narrow base and sharp peak is representative of a low-contrast image and therefore of a homogeneous sample. Conversely, an asymmetric histogram with a wide base and flatter peak or several modes is representative of a more contrasted image, i.e. a heterogeneous sample. Four metrics are typically used [160] to characterize distribution on the image histogram: mean; variance, describing variation about the mean; skew, which measures asymmetrical tailing, a positive skew indicating tailing towards higher values and a negative skew tailing toward lower values; and kurtosis, or pointiness, which gives information about the shape of the histogram peak. As the kurtosis of a Gaussian distribution is 3, then kurtosis >3 indicates sharper peaks with a long tail whereas kurtosis <3 describes flatter peaks with a smaller tail. Histogram analysis of images extracted in univariate or multivariate fashion has frequently been used to assess sample homogeneity [76,101,104,160,161].

#### 3.3.3. Image binarization

Image binarization aims at separating the object of interest (foreground) from background. In some cases the patterns to be segmented have nearly the same pixel intensities, meaning that the histogram has a single mode. Thresholds can be set between histogram modes so that pixels falling between two threshold limits belong to the object of interest and those outside are rejected as background. The threshold can be user-defined, as in most publications on the pharmaceutical applications of CI. The main disadvantage is that fixing the threshold is subjective, meaning that threshold values differ between users. Automated methods such as Otsu's [162] may be an alternative solution. The problem of image binarization by histogram analysis is that many images do not have clear separate modes or they may have a mode which does not correspond to a distinct structure in the image. Otherwise, most of the clustering techniques discussed in Section 3.2.2.2 (KM, Gaussian mixtures, ANN, etc.) can also be applied to classify pixels into foreground and background so as to produce a binary image.

Binary images are also known as Boolean images because they consist of pixels having only two possible values: 1 (mostly associated with foreground) and 0 (mostly associated with background). However, a binary image is rarely perfect and the pixels may be misclassified. For this reason, binarization has to be refined, e.g. by introducing morphological operations [159]. Boolean procedures generate a new image by combining different binary images of the same scene using Boolean AND/OR procedures. Morphological operations act by setting pixels to 1 or 0 according to simple rules based on the value of neighboring pixels. For example, they remove individual pixels, and narrow or widen edges. A major difficulty is separating contiguous features in a binary image.

Once an image has been properly segmented, information such as mean particle size and S.D. can be extracted. The percentage area covered by the compounds can also be correlated with the true content of the sample [50,160,163].

## 4. Pharmaceutical applications

### 4.1. Sample preparation and measurement

For measurement to be as representative of a sample's chemical composition as possible, all physical effects such as superficial roughness must be minimized. LaPlant has provided detailed advice on the preparation of powders and tablets [164]. Trustworthy images from powders such as raw granulation or capsule fill are difficult to obtain. A press can be used to generate a kind of pill but compression may compromise the test sample. Tablets are best milled with a bevel-edged blade to remove the coating and flatten the surface (Fig. 12). However, depending on tablet hardness and friability, this may remove or displace some particles from the surface for analysis.

It is difficult to be categorical about choosing an objective. Is it better to use high spatial resolution to detect fine patterns or low spatial resolution to cover a larger area of analysis? Best advice is to take time first to map the whole surface, then focus on a particular region at higher spatial/spectral resolution if minor compounds need resolving [39]. This naturally depends on the particle sizes of the tablet constituents and the spectroscopic technique used. Ma and Anderson [165] tested objectives with different pixel sizes: 67.1 μm/pixel, 45.5 μm/pixel and 21.5 μm/pixel on a binary mix-

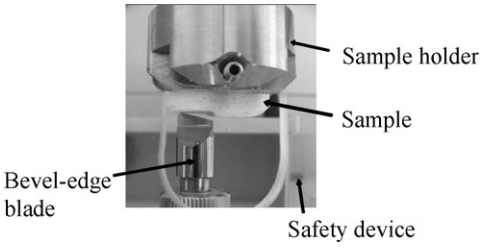


**Fig. 12.** Preparing the tablet surface for analysis: milling with a bevel-edge blade.

ture blend containing 80% lactose, median particle size 100 μm, and 20% salicylic acid (SA), median particle size 30 μm. They reported that "higher magnification levels did not provide additional relevant information about the surface features" even when the spatial resolution of the higher magnification objective was lower than the median particle size. The 67.1 μm/pixel objective was thus preferable.

### 4.2. Chemical distribution

Several studies have addressed the extraction of chemical distribution maps from solid dosage forms, often comparing processing or acquisition modes in order to evaluate the pros and cons of each. In NIR reflectance mode, single wavelength images, RGB reconstruction, and PCA analysis all reveal tablet compound distribution maps [166–168], even successfully extracting the distribution maps of compounds in a timed-release granule [168]. Such tablets form complex systems in several layers, each having its own functionality in active substance release. Thus NIR-CI is an effective method of controlling compound distribution, which is an important factor in the dissolution behavior of the medicine.

Comparison between micro- and macro-ATR accessories for FT-MIR imaging of pharmaceutical formulations showed the main difference to reside in the spatial resolution that could be achieved. Macro-ATR enabled an area of ~1 $mm^2$ to be scanned, giving an overall view of the tablet surface, whereas micro-ATR could analyze an area of 250 $\mu m^2$, detecting compounds at concentrations as low as 0.5% [90].

Thanks to its specificity and fine spatial resolution, Raman spectroscopy in combination with advanced chemometric tools can identify compounds present at concentrations as low as 0.2% [114]. In a comparison between NIR-CI and Raman imaging at the same magnification levels, Šašic [98] found that Raman imaging combined with multivariate analysis detected all five compounds in a tablet whereas NIR imaging failed to extract the two minor ones. In a subsequent study by Šašic [169] comparing global Raman imaging and NIR chemical mapping for detecting the chemical composition of granules, both techniques showed the granules to be mainly a mixture of API and mannitol whereas some granules contained pure compounds. In the case of this formulation, compound visualization was better with Raman imaging.

Many studies have also shown that classic multivariate analysis, e.g. PCA, OLS or direct classical least squares (DCLS), MCR and CCA, combined with Raman imaging, provides reliable information about tablet compound distribution [65,93,140,170]. More recent tools such as 2D correlation [154,155] also accurately revealed the localization of chemicals inside tablets containing up to four components.

Fig. 13 gives an example of a micro-RGB image of a tablet. The low-dosage tablet was analyzed using a Raman spectrometer equipped with a line detector. Images of the Raman shift at 269 $cm^{-1}$ for the API, 427 $cm^{-1}$ for cellulose, and 347 $cm^{-1}$ for lactose were assigned to the red, green, and blue channels of the RGB image, respectively. The spectra localized in the domain specific to each compound accurately matched the reference spectra (Fig. 13, right plots). Correlation coefficients to the reference spectra were 0.98 for the API, 0.91 for cellulose, and 0.96 for lactose, clearly confirming the compounds' identities.

### 4.3. Blending

During pharmaceutical production, blend homogeneity is essential if quality attributes, such as dissolution behavior, are to be maintained, and if all tablets are to contain the correct amount of API. The classic method of testing blend content uniformity is to remove an amount of powder using a sample thief and check its content by high performance liquid chromatography (HPLC) or UV–vis spectroscopy. CI is now being studied as an alternative approach.

Classical NIR spectroscopy and CI were investigated as potential noninvasive methods with three binary mixtures composed of Fast-Flo lactose and variable amounts of SA [171]. NIR analysis was performed at different positions in the blender through six sapphire windows. Six images were acquired at different wavelengths spanning the lactose and SA peaks. The blender lids were removed to allow access to the top surface of the powder for imaging, directly in the blender (i.e. without sample removal). The reference method using classic sampling combined with UV spectroscopy was run in parallel as an accuracy control. Cluster analysis and moving block S.D. along wavelength values (spectral space) were used to analyze the classical NIR spectra. The moving block S.D. procedure was applied in multidimensional space. The study showed agreement between the NIR spectroscopic measurements and the reference methods. NIR imaging sampled a larger area (~15 $cm^2$) but only the upper layer in this experimental configuration.

NIR-CI and classic NIR spectroscopic analysis were compared in a study of tablets produced after different mixing times [172]. Commercial tablets were also used for further comparison with experimental tablets. Constituent homogeneity in the final tablets was assessed by histogram analysis of univariate images or images extracted by PLS. NIR imaging was better than NIR spectroscopy at detecting slight problems of homogeneity. The advantage of NIR spectroscopy is that it can investigate microdomains within tablets whereas classical integrating NIR spectroscopy tends to overlook slight heterogeneity.

A prototype fiber bundle-CI system for on-line inspection of dry blend homogeneity has been recently, and briefly, described [173]. The first steps in its codevelopment by Spectral Dimensions (now Malvern) and Pfizer were described at www.in-PharmaTechnologist.com in 2005 [174]. This kind of equipment would allow on-line blend characterization by CI.

The wavelet transform has been successfully combined with PCA to determine the texture of pharmaceutical tablets and characterize the density and size of differently dosed API (paracetamol) [175].

Fig. 14 demonstrates the ability of NIR-CI to detect nonuniformity in a low-dosage tablet containing 3% API. After SNV normalization and second derivative transformation, API, cellulose, and lactose were mapped by selecting images at specific wavelengths (2140 nm, 2000 nm, and 1900 nm), followed by RGB reconstitution. The white circles clearly show the aggregates of API.

### 4.4. Content uniformity

Several studies have addressed the potential of CI to calculate compound concentrations within samples. Jovanovic et al. [104] used binary protein–sugar mixtures (lysozyme–trehalose), at concentrations of 0%, 10%, 50%, 80%, and 100% (w/w) lysozyme, to compare methods for content. They created quantitative models based on correlation coefficient and PLS regression to relate the contrast directly to the concentration differences of each constituent. The method was then applied to dried and freeze-dried samples containing 50% lysozyme and 50% trehalose. PLS-DA regression was better at predicting concentrations in the calibration mixtures, but predictions with dried and freeze-dried samples were worse than expected, probably due to extra spectral variations in the predicted sample compared to the spectral library. CI proved a useful tool for confirming mixture homogeneity.

The present authors used NIR imaging to predict API, cellulose and lactose in low-dosage pharmaceutical tablets [176]. API content was varied from 0% to 10% in 1% increments. Four tablets were

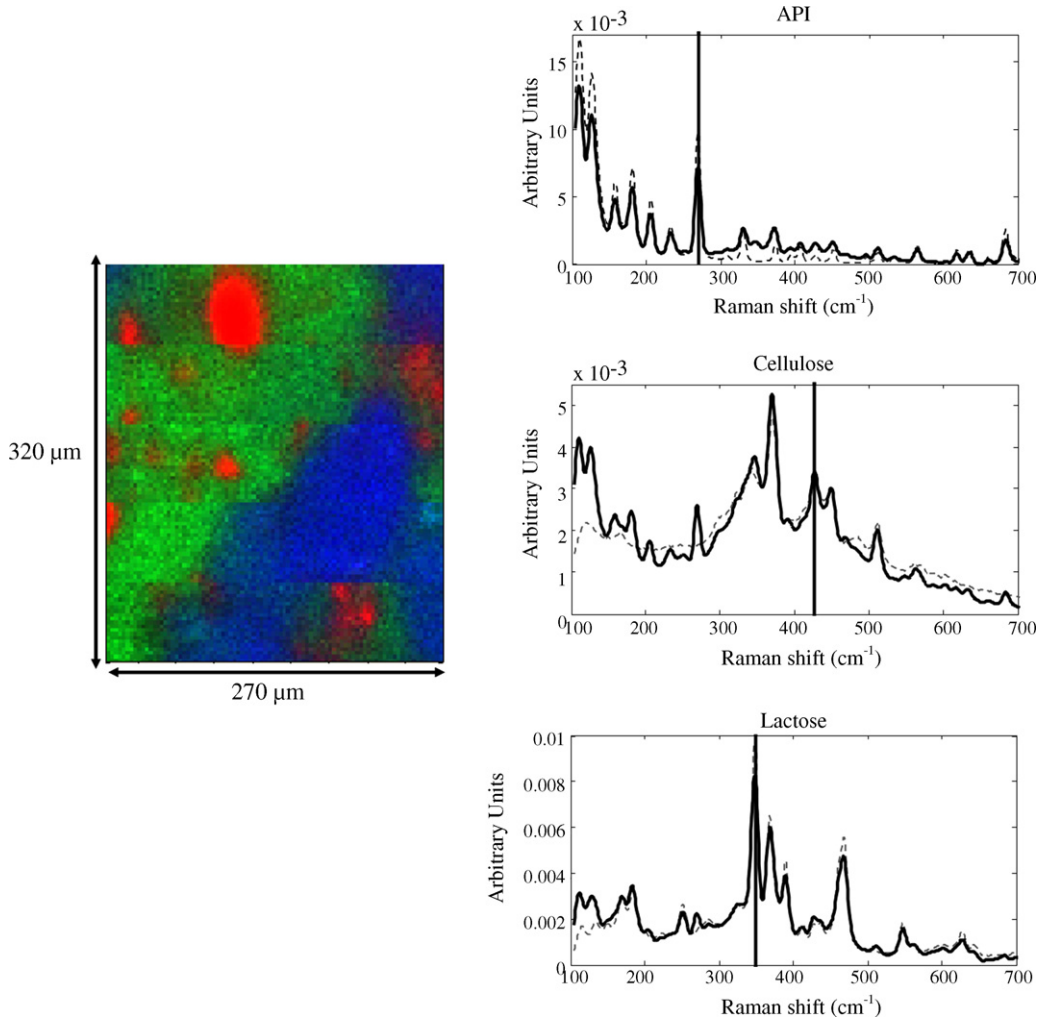


**Fig. 13.** Micro-scale red–green–blue (RGB) image of a tablet. The data were acquired with a Raman spectrometer equipped with a line detector. Red depicts active pharmaceutical ingredient (API: white in the black-and-white version), green cellulose (gray in the black-and white-version), and blue lactose (black in the black-and-white version). The spectra on the right were extracted from the data cube. API spectra were extracted from the red area, cellulose spectra from the green area and lactose spectra from the blue area. The spectra extracted from the data cube match the reference spectra (dotted line). Vertical lines: the wavelengths chosen to build the RGB image (269 $cm^{-1}$ for the API, 427 $cm^{-1}$ for cellulose, and 347 $cm^{-1}$ for lactose). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

measured from each batch. A PLS model was built from the average spectra and used to predict each pixel spectra. The concentrations of API, cellulose, and lactose were calculated per data cube by averaging each pixel prediction. DCLS based on pure spectra was also tested as a method of prediction. Two preprocessing techniques were investigated: normalized spectra and normalized spectra followed by second derivative transformation. PLS was able to extract accurate predictions of API, cellulose, and lactose with both preprocessing treatments. DCLS worked well for the prediction of API with second derivative spectra but not for the prediction of cellulose and lactose. Hence, as with classical NIR spectra, PLS regression is the method of choice for content prediction.

High-throughput NIR has recently been proposed for API content [177]. Instead of measuring calibration and test samples separately, a wide FOV (59.5 mm × 47.5 mm) was used and tablets were measured simultaneously. A single image at 1600 nm (API peak) featured a trend correlated to the API concentration and this wavelength was therefore chosen for controlling content uniformity. A calibration curve based on NIR intensities at 1600 nm was computed from calibration samples enabling unknown samples to be subsequently predicted. Analysis was rapid and reliable in predicting API in comparison with UV. Analyzing calibration and test samples simultaneously overcomes the problem of day-to-day repeatability and instrument stability (e.g. waning lamp power). In addition, if a single wavelength suffices, then measurement is fast and data volume reduced. The authors estimated that by widening the FOV, it might be possible to analyze up to 1500 tablets simultaneously, resulting in 100% control.

Raman imaging has also been used to predict API content. Results with Raman point mapping and MCR-ALS in semi-quantitative analysis showed errors up to 8% in predicting API or excipient concentration [178].

### 4.5. Polymorphism

Lattice of a given crystal may vary, or a molecule may bind differently with water or solvent (hydrate and solvate) leading to different physical forms. It is important to detect and quantify polymorphic forms because they can influence a tablet's physical properties. Detection is more readily achieved with Raman imaging whereas NIR or MIR is more appropriate for differentiating hydrates and solvates [39].

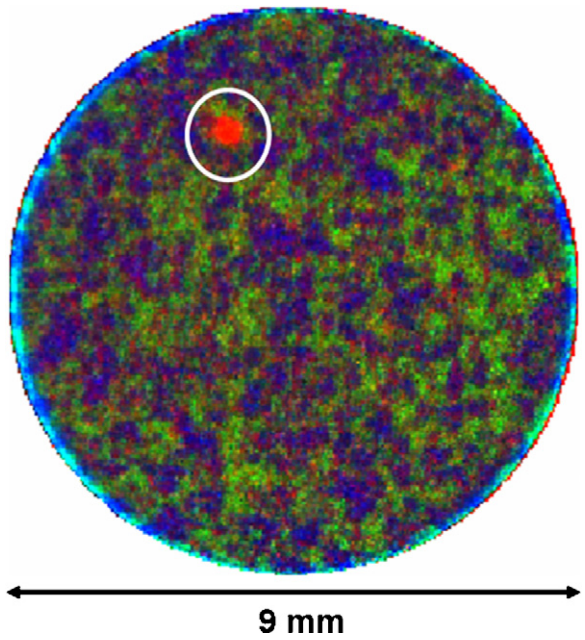


**Fig. 14.** Macro-scale red–green–blue (RGB) image of a tablet containing 3% API. The tablet was scanned by an NIR imaging system. The objective gives a spatial resolution of 40 μm/pixel. Red depicts API (white in the black-and-white version), green cellulose (gray in the black-and-white version), and blue lactose (black in the black-and-white version). An API agglomerate is clearly visible (white circle). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

Clark et al. described several examples of Raman mapping applications for polymorph detection [39]. Two dealt with the localization of different API forms in raw API material, detecting fine impurities at concentrations as low as 1% (w/w). Others involved the visualization of different API forms in a formulation blend. The authors claimed that API impurities could be detected at concentrations of 0.025%. They also showed that Raman imaging could identify polymorphic changes due to product manufacturing such as hot melt extrusion or coating. In addition they used NIR imaging to measure different levels of drug hydration due to varying process blending or granulation methods.

Henson and Zhang [179] reported detecting 0.05% (w/w) polymorphic impurity in a tablet matrix using Raman imaging [158]. They applied directed agglomeration clustering (DAC) to distinguish between two polymorphs. A set of binary mixtures with differing contents (0%, 25%, 50%, 75%, and 100% of form 1) was constructed; predicted concentrations proved a close match. Raman microscopy has also been combined with atomic force microscopy (AFM) to investigate surface domains of amorphous sorbitol [180]. The two methods were complementary, AFM being more sensitive to surface differences and Raman providing lateral and depth information about amorphous material.

Chan et al. [40,41] showed how transmission FTIR could be deployed in monitoring polymorphic changes under a controlled environment. They successfully visualized the transition of nifedipine from the β form to the α form at 38 °C and 80% humidity for 4 h [40,41]. They also studied four binary mixtures of nifedipine and nitrendipine films at 40 °C and 90% humidity, showing both compounds to be stable in the glassy phase at equimolar concentrations. The advantage of this approach is that the samples were analyzed simultaneously thanks to a macro chamber and wide array detector (128 × 128 pixels) which made robust comparison of the mixtures possible.

### 4.6. Dissolution and drug delivery

Classic dissolution tests and UV spectroscopy can determine how much drug has been released over time. However, these techniques give no insights into the mechanism of drug dissolution, nor allow study of any crystallization processes that may occur. Spatial information has the potential to overcome these limitations, hence the application of FTIR imaging to the study of drug release. Two laboratories in particular have used CI to investigate dissolution.

Jack Koenig's laboratory at Case Western Reserve University, Cleveland, studied the diffusion of nicotine in ethanol/deuterium oxide ($D_2O$) mixtures into an ethylene–vinyl acetate (EVA) copolymer membrane as part of an investigation into transdermal drug delivery systems [181]. Their aim was to determine how different dilutions of ethanol (0–100% in 20% increments) affected diffusion. The measurements were done in transmission and they used peak integration to monitor swelling of the EVA membrane over time and spatial compound distribution. They showed that nicotine diffused into the EVA membrane ahead of the solvent at ethanol concentrations ≤60%. When ethanol was present at 80%, $D_2O$ led the diffusion front, but at the 100% concentration ethanol was first. They observed an exponential trend for an increase in swelling rate with increasing ethanol in the solvent. Thus adjusting the solvent composition could speed or slow the nicotine diffusion rate. Koenig's group then used FTIR imaging to study the dissolution in $D_2O$ of testosterone suspended in polyethylene oxide (PEO) matrix [182]. They measured four binary mixtures containing different amounts of drug (10%, 20%, 30%, and 40%), drawing the following conclusions about the role of drug loading from the spatial information obtained: in mixtures containing 10% and 20% testosterone, polymer dissolution controlled the drug release, whereas at 30% and 40% dissolution of the drug dictated its release rate. Spectral variations also revealed localized changes in molecular structures.

Kazarian and Chan at Imperial College London drew on the Koenig group's work to study the release of different polymer/drug formulations into water: polyethylene glycol (PEG)/ibuprofen, PEG/sodium benzoate, and methyl-β-cyclodextrin (MBCD)/ibuprofen [183]. Due to the strong absorption of water in the MIR range, they acquired spectral images using a macro ATR accessory. They observed that the poorly soluble drug, ibuprofen, crystallized during the addition of water whereas the highly soluble sodium benzoate did not. The initial conclusion was that drug solubility was essential and that crystallization should be avoided because of its impact on bioavailability. Experience then showed that no crystallization occurred with the MBCD/ibuprofen mixture, i.e. that choosing an appropriate polymer improved release behavior. Kazarian et al. since continued their work on the dissolution of synthetic binary tablets in water using macro-ATR [184–187]. They designed a cell in which drug mixture was compacted on the surface of a diamond crystal and flushed with water to prevent saturation [184]. Study of the dissolution of a mixture containing 20% caffeine and 80% hydroxypropylmethylcellulose (HPMC) validated the device for dissolution monitoring. Subsequently [185], they showed that a ZnSe crystal can also be used for *in situ* dissolution even if the compaction force has to be lowered: its advantage is a wider FOV that visualizes the entire tablet surface; they also introduced PLS calibration to monitor concentration of HPMC release over time. In their next study [186], using UV to assay the liquid collected at the cell outflow for the mixture of HPMC and niacinamide, they showed that the concentrations predicted by PLS regression closely matched the UV values and enabled distribution maps to be extracted over time for HPMC, drug, and water. Their most recent study [187] addressed the dissolution of tablets containing poorly soluble drug (diclofenac) and HPMC. Medium pH was varied, and images

were extracted using DCLS. Drug release profiles were found not to be pH-dependent. Precipitation was also detected and roughly quantified.

The above studies confirm the utility of FTIR imaging for elucidating drug release mechanisms. However, all the test samples considered were synthetic binary mixtures and real final pharmaceutical solids often form a much more complex matrix.

### 4.7. Process understanding, troubleshooting, and product design

CI now plays an important role in new formulation development and process troubleshooting. This is because spatial compound distribution is an important factor in drug behavior, with particular respect to dissolution profile, stability, and bioavailability.

Reich [53] described two applications of FTIR imaging: control of film coat uniformity on tablets, and chemical mapping of the distribution of a model protein in polymer matrix tablets (poly[DL-lactide-*co*-glycolide] (PLGA)) immediately after processing and during the release phase. The results showed that curing the tablets produced a more uniform coating, and that after a period under *in vitro* release conditions, the protein appeared on the surface, indicating absorption effects.

Kazarian et al. have studied the effects of process parameters on various solid dosage forms in four publications. Using an experimental set-up that visualized water vapor sorption into the different domains of a PEG–griseofulvin mixture under different relative humidities (0.5–90% at 25 °C), they showed that the polymer mainly absorbed the water, especially above 70%, while spatial drug distribution remained unchanged [42]. Next, using a cell plate allowing the simultaneous monitoring of 100 polymer/drug formulations under identical conditions (relative humidity 67% for 30 min), they studied several PEG/nifedipine mixtures containing varying drug concentrations [91]. Formulations containing more drug showed lower water sorption. In another important study for the optimization of tablet production, they showed how tablet density increases with the addition of lubricant [57]. Most recently they have addressed the effects of moisture and pressure on tablet compaction in a model drug composed of HPMC and ibuprofen, showing how tablet density increases with relative humidity [188].

MIR hyperspectral imaging has been used for process troubleshooting on actual pharmaceutical formulations [10,189], in particular to resolve contamination and dissolution issues. Images at specific wavelengths and PCA were used to extract the compound distribution maps. The contaminant was revealed as a mixture of indigo carmine and excipients, with a high residual moisture content [10,189], while the dissolution issue was explained by an excess of magnesium stearate on the tablet surface preventing water action.

Hammond and Clarke [11] described four applications of NIR point mapping to solve process issues: (1) NIR-CI investigation of differing flow properties within a batch revealed nonuniform distribution of lubricant as the root cause of poor flow; (2) NIR mapping revealed areas with excess inorganic filler as the cause of tablets sticking to the press punches; (3) tablet chipping after a change in sugar supplier was explained by better mixing of the new sugar with the inorganic filler, resulting in weaker tablets more prone to chipping; and (4) lubricant was revealed as the tablet excipient that absorbed water.

Clarke [50] investigated two different batches. The first had a problem of processability. In the bad samples, NIR imaging (generated with an FPA detector and processed by PCA analysis and RGB reconstruction) revealed larger domain sizes of polymer components. In the second, NIR images generated by line mapping explained how tablet dissolution properties depended on roller compaction force: the cluster size of disintegrant material increased with compression force.

Our own laboratory used CI to assess homogeneity in two intermediates produced using different process parameters (screw speed and API particle size) [161]. With the first intermediates, sampled just after extrusion, screw speed was found to influence the roughness of the extrudates as well as their homogeneity. With the second intermediates (uncoated cores), CLS analysis revealed that higher screw speed produced finer API granulates within the cores.

NIR-CI has been used to assess the difference of tartaric acid particle size between tablet tops and bottoms [190]. Three different batches were produced using different API particle sizes. A statistically significant top/bottom difference in particle size was found in tablets produced with larger API particles. API segregation was found to occur during the compression phase with the larger particle size. Thus segregation can be minimized using fine-particle API. NIR-CI has also shown that larger API particle sizes result in blends with larger API agglomerates [163].

NIR-CI can also detect variation due to differences in tablet compression force, i.e. it is an analytical tool for explaining physical variations [191]. In another recent study, it was also used to measure the distribution of density and tabletting force in pharmaceutical tablets containing different amount of magnesium stearate (MgS). The tablet density was less uniform with low content of MgS [192].

### 4.8. Particle size

Doub et al. [193] used Raman widefield imaging to characterize aqueous suspension nasal spray formulations. In a preliminary validation step, they evaluated the size of polystyrene microspheres. The results showed good agreement with nominal values and no statistically significant differences between mean particle sizes. In actual samples of pure micronized drug substance, image analysis of particle sizes showed good agreement with laser light scattering (LLS) in one half of the batches but not the other half. The ability of Raman imaging to discriminate between API and excipients in nasal spray solution was also demonstrated. But particle sizes computed with Raman imaging differed from the LLS values. The authors suggested further work (sample preparation, scanning more particles, and studying formulations with different API particle sizes) to improve the results.

### 4.9. Counterfeit and identification

Counterfeit is now a serious problem as it can damage health by supplying inappropriate substances or products devoid of API. CI is a useful weapon in rapid counterfeit detection. Our own laboratory analyzed tablets from a suspect lot and original product using a wide FOV, processing both data cubes simultaneously by PCA (Fig. 15). The score images of the second principal component (Fig. 15b) revealed great chemical differences between the original and the product under investigation, which was therefore counterfeit.

NIR-CI is an effective tool for discriminating between fake and real medicines [49]: samples were analyzed in the same FOV with the real medicine as reference; PCA score plots revealed samples containing no API or differing in chemical composition. MIA of NIR images can also detect small differences in tablet uniformity that discriminate between counterfeit and real product [194]. FT-MIR imaging was combined with desorption electrospray ionization linear ion trap mass spectrometry (DESI-MS) to identify chemical species in a counterfeit antimalarial, finding no API present [195].

Websites offering cutprice pharmaceuticals are increasingly common. The quality of such products may be inferior to that of the genuine article and requires checking. NIR-CI was used to com-

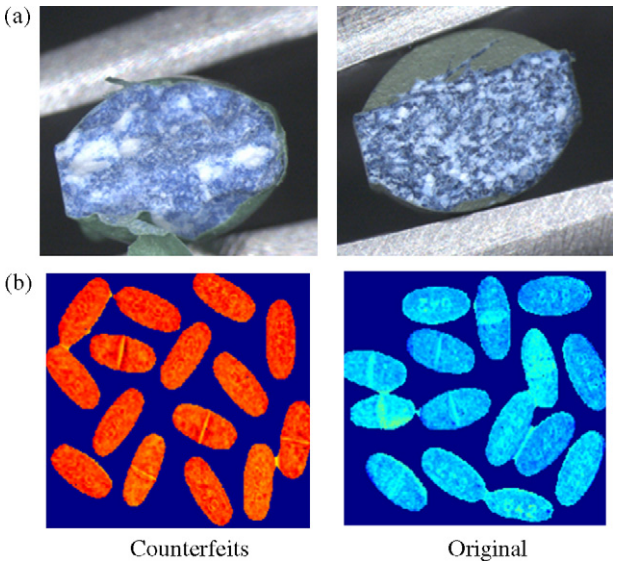


**Fig. 15.** (a) Photograph of a suspected counterfeit (left) and the original product (right). The suspected counterfeit shows nonuniformity. (b) Score images for the second principal component. PCA reveals clearly different composition between the two tablets.

pare generic tablets imported from Mexico, India, Thailand, Brazil, and Canada via the Internet with the US manufacturer's tablets [196,197]. The Canadian tablets were similar in API blend uniformity to the US product and hence presumably of the same quality. However, those from the other countries differed in API distribution, and in particular in their API agglomerates, suggesting inferior quality. NIR-CI has become an analytical tool for assessing Internet drug quality, with several Internet products failing the USP monograph for quality attributes due to differences in chemical composition and compound distribution [105].

### *4.10. Tablet imaging through blister packaging*

Multispectral imaging has been used to evaluate water content and identify thousands of individual tablets through blister packaging [198]. Water content was analyzed by pricking the blister and exposing the tablets to water vapor for 24 h, with six tablets being sampled hourly for calibration. Calibration was obtained by measuring 43 tablets at 0.5 m. The standard error of estimate was 0.06%. In the identification part of the study, some 1300 tablets could be visualized simultaneously, and SA could still be differentiated from acetylsalicylic acid (ASA) by PCA analysis (as confirmed in Ref. [199]).

## 5. Conclusions

Due to their complementarity and specificity, IR, NIR, and Raman spectroscopy have been widely used to characterize chemical species. The possibility of acquiring spatially located spectra adds a valuable dimension to the analysis of matter. Since the first appearance of chemical maps, a wide range of devices has been developed allowing various modes of acquisition and a range of spectral and spatial resolutions. Such has been the technical advance that a full data cube can now be acquired within minutes. Users can devise their own menu of preferred spectral/spatial information and set up the device accordingly for faster acquisition. The parallel emergence of new software (algorithms, computational tools, and chemometric and image-processing programs) has optimized the processing of user-relevant information. Collaboration between imaging and chemical experts will only improve data handling further.

By combining spectral and spatial information, CI provides significant insights into pharmaceutical samples. There is now incontrovertible evidence that API and excipient distribution is an important quality feature because of its potential impact on tablet behavior. That accounts for the great interest in CI for pharmaceutical applications such as blend monitoring, polymorph characterization, and process monitoring. However, no fully on-line applications of CI for 100% quality control are as yet operational. The NIR pushbroom imagers that have been launched onto the market are fast, but still not fast enough. On-line applications promise to be a huge growth field for CI in the near future.

## References

[1] Y. Roggo, P. Chalus, L. Maurer, C. Lema-Martinez, A. Edmond, N. Jent, J. Pharm. Biomed. Anal. 44 (2007) 683–700.
[2] S. Wartewig, R.H.H. Neubert, Adv. Drug Deliv. Rev. 57 (2005) 1144–1170.
[3] R. Barer, A.R.H. Cole, Nature 163 (1949) 198–201.
[4] M.A. Harthcock, S.C. Atkin, Appl. Spectrosc. 42 (1988) 449–455.
[5] P. Treado, I.W. Levin, E.N. Lewis, Appl. Spectrosc. 48 (1994) 607–615.
[6] W.H.A.M. van den Broek, E.P.P.A. Derks, E.W. van de Ven, D. Wienke, P. Geladi, L.M.C. Buydens, Chemometr. Intell. Lab. Syst. 35 (1996) 187–197.
[7] J.R. Mansfield, L.M. McIntosh, A.N. Crowson, H.H. Mantsch, M. Jackson, Appl. Spectrosc. 53 (1999) 1323–1330.
[8] A.A. Gowen, C.P. O'Donnell, P.J. Cullen, G. Downey, J.M. Frias, Trends Food Sci. Technol. 18 (2007) 590–598.
[9] http://www.fda.gov/cder/guidance/6419fnl.pdf, 2007.
[10] Y. Roggo, N. Jent, A. Edmond, P. Chalus, M. Ulmschneider, Eur. J. Pharm. Biopharm. 61 (2005) 100–110.
[11] S.V. Hammond, F.C. Clarke, in: J.M. Chalmers, P.R. Griffiths (Eds.), Handbook of Vibrational Spectroscopy, vol. 2, John Wiley & Sons, London, 2002, pp. 1405–1418.
[12] Y. Roggo, M. Ulmschneider, in: S.C. Gad (Ed.), Pharmaceutical Manufacturing Handbook: Regulations and Quality, Wiley, 2008.
[13] D. Steele, in: J.M. Chalmers, P.R. Griffiths (Eds.), Handbook of Vibrational Spectroscopy, vol. 1, John Wiley & Sons, London, 2002, pp. 44–70.
[14] R.W. Hannah, in: J.M. Chalmers, P.R. Griffiths (Eds.), Handbook of Vibrational Spectroscopy, vol. 2, John Wiley & Sons, London, 2002, pp. 933–952.
[15] G. Keresztury, in: J.M. Chalmers, P.R. Griffiths (Eds.), Handbook of Vibrational Spectroscopy, vol. 1, John Wiley & Sons, London, 2002, pp. 71–87.
[16] P.R. Griffiths, in: J.M. Chalmers, P.R. Griffiths (Eds.), Handbook of Vibrational Spectroscopy, vol. 1, John Wiley & Sons, London, 2002, pp. 33–43.
[17] D. Bertrand, V. Baeten, in: D. Bertrand, E. Dufour (Eds.), La spectroscopie infrarouge et ses applications analytiques, Tech & doc, Paris, 2006, pp. 247–301.
[18] C.J. Manning, in: J.M. Chalmers, P.R. Griffiths (Eds.), Handbook of Vibrational Spectroscopy, vol. 2, John Wiley & Sons, London, 2002, pp. 283–297.
[19] L.H. Kidder, A.S. Haka, E.N. Lewis, in: P.R. Griffiths (Ed.), Handbook of Vibrational Spectroscopy, vol. 2, John Wiley & Sons, London, 2002, pp. 1386–1404.
[20] B. Schrader, in: J.M. Chalmers, P.R. Griffiths (Eds.), Handbook of Vibrational Spectroscopy, vol. 2, John Wiley & Sons, London, 2002, pp. 1289–1301.
[21] D.L. Wetzel, A.J. Eilert, J.A. Sweat, in: J.M. Chalmers, P.R. Griffiths (Eds.), Handbook of Vibrational Spectroscopy, vol. 1, John Wiley & Sons, London, 2002, pp. 436–452.
[22] N. Gat, Proc. SPIE 4056 (2000) 50–64.
[23] M.J. Pelletier, in: J.M. Chalmers, P.R. Griffiths (Eds.), Handbook of Vibrational Spectroscopy, vol. 1, John Wiley & Sons, London, 2002, pp. 467–481.
[24] F.A. DeThomas, P.J. Brimmer, in: J.M. Chalmers, P.R. Griffiths (Eds.), Handbook of Vibrational Spectroscopy, vol. 1, John Wiley & Sons, London, 2002, pp. 383–392.
[25] M. Ford, in: J.M. Chalmers, P.R. Griffiths (Eds.), Handbook of Vibrational Spectroscopy, vol. 1, John Wiley & Sons, London, 2002, pp. 249–253.
[26] P. Colarusso, L.H. Kidder, I.W. Levin, J.C. Fraser, J.F. Arens, E.N. Lewis, Appl. Spectrosc. 52 (1998) 106A–120A.
[27] C.D. Tran, Appl. Spectrosc. Rev. 38 (2003) 133–153.
[28] E.N. Lewis, P.J. Treado, R.C. Reeder, G.M. Story, A.E. Dowrey, C. Marcott, I.W. Levin, Anal. Chem. 67 (1995) 3377–3381.
[29] M.D. Schaeberle, H.R. Morris, J.F. Turner, P. Treado, Anal. Chem. (1999) 175A–181A.
[30] A.J. Sommer, in: J.M. Chalmers, P.R. Griffiths (Eds.), Handbook of Vibrational Spectroscopy, vol. 2, John Wiley & Sons, London, 2002, pp. 1369–1385.
[31] J.M. Chalmers, P.R. Griffiths, in: D.E. Pivonka, J.M. Chalmers, P.R. Griffiths (Eds.), Applications of Vibrational Spectroscopy in Pharmaceutical Research and Development, John Wiley & Sons, Chichester, 2007, pp. 19–49.
[32] S. Šašic, D.A. Clark, Appl. Spectrosc. 60 (2006) 494–502.
[33] B.M. Patterson, G.J. Havrilla, Appl. Spectrosc. 60 (2006) 1256–1266.

[34] P.J. Treado, M.P. Nelson, in: J.M. Chalmers, P.R. Griffiths (Eds.), Handbook of Vibrational Spectroscopy, vol. 2, John Wiley & Sons, London, 2002, pp. 1429–1459.
[35] R. Bhargava, I.W. Levin, Anal. Chem. 73 (2001) 5157–5167.
[36] R. Bhargava, T. Ribar, J.H. Koenig, Appl. Spectrosc. 53 (1999) 1313–1322.
[37] R.L. McCreery, Raman Spectroscopy for Chemical Analysis, John Wiley & Sons, 2000, pp. 293–332.
[38] S. Schlucker, M.D. Schaeberle, S.W. Huffman, I.W. Lewin, Anal. Chem. 75 (2003) 4312–4318.
[39] D. Clark, M. Henson, F. Laplant, S. Šašic, in: D.E. Pivonka, J.M. Chalmers, P.R. Griffiths (Eds.), Applications of Vibrational Spectroscopy in Pharmaceutical Research and Development, John Wiley & Sons, London, 2007, pp. 309–335.
[40] K.L.A. Chan, S.G. Kazarian, D. Vassou, V. Gionis, G.D. Chryssikos, Vib. Spectrosc. 43 (2006) 221–226.
[41] K.L.A. Chan, S.G. Kazarian, Vib. Spectrosc. 42 (2006) 130–134.
[42] K.L.A. Chan, S.G. Kazarian, Vib. Spectrosc. 35 (2004) 45–49.
[43] P. Lasch, D. Naumann, Biochim. Biophys. Acta (BBA)-Biomembr. 1758 (2006), 841-829.
[44] C. Krafft, L. Shapoval, S.B. Sobottka, G. Schackert, R. Salzer, Biochim. Biophys. Acta (BBA)-Biomembr. 1756 (2006) 883–891.
[45] C. Beleites, G. Steiner, M.G. Sowa, R. Baumgartner, S. Sobottka, G. Schackert, R. Salzer, Vib. Spectrosc. 38 (2005) 143–149.
[46] J.R. Mansfield, M.G. Sowa, C. Majzels, C. Collins, E. Cloutis, H.H. Mantsch, Vib. Spectrosc. 19 (1999) 33–45.
[47] J.A. Fernandez Pierna, V. Baeten, A. Michotte Renier, R.P. Cogdill, P. Dardenne, J. Chemometr. 18 (2004) 341–349.
[48] J.C. Noordam, W.H.A.M. van den Broek, P. Geladi, L.M.C. Buydens, Chemometr. Intell. Lab. Syst. 75 (2005) 115–126.
[49] J. Dubois, J.-C. Wolff, J.K. Warrack, J. Schoppelrei, E.N. Lewis, Spectroscopy (2007).
[50] F. Clarke, Vib. Spectrosc. 34 (2004) 25–35.
[51] F.M. Mirabella, in: J.M. Chalmers, P.R. Griffiths (Eds.), Handbook of Vibrational Spectroscopy, vol. 2, John Wiley & Sons, London, 2002, pp. 1091–1102.
[52] M. Meurens, in: D. Bertrand, E. Dufour (Eds.), La spectroscopie infrarouge et ses applications analytiques, Tech & doc, Paris, 2006, pp. 233–245.
[53] G. Reich, Pharmind 64 (2002) 870–874.
[54] C.M. Snively, J. Lauterbach, Spectroscopy 17 (2002) 26–33.
[55] S. Kazarian, Anal. Bioanal. Chem. 388 (2007) 529–532.
[56] K.L.A. Chan, S.G. Kazarian, Lab on a Chip-Miniaturisation Chem. Biol. 6 (2006) 864–870.
[57] K.L.A. Chan, N. Elkhider, S.G. Kazarian, Chem. Eng. Res. Des. 83 (2005) 1303–1310.
[58] P. Geladi, J. Burger, T. Lestander, Chemometr. Intell. Lab. Syst. 72 (2004) 209–217.
[59] J. Burger, P. Geladi, J. Chemometr. 19 (2005) 355–363.
[60] Z. Liu, H. Yu, J.F. MacGregor, J. Chemometr. 21 (2007) 88–95.
[61] R.L. McCreery, Raman Spectroscopy for Chemical Analysis, John Wiley & Sons, 2000, pp. 251–292.
[62] Chan, Kazarian, Appl. Spectrosc. Rev. 57 (2003) 381–389.
[63] A.J. Sommer, L.G. Tisinger, C. Marcott, G.M. Story, Appl. Spectrosc. 55 (2001) 252–256.
[64] L.L. Lewis, A.J. Sommer, Appl. Spectrosc. 54 (2000) 324–330.
[65] D. Clark, S. Šašic, Cytometry 69 (2006) 815–824.
[66] L. Duponchel, P. Milanfar, C. Ruckebusch, J.-P. Huvenne, Anal. Chim. Acta 607 (2008) 166–175.
[67] S.J. Hudak, K. Haber, G. Sando, L.H. Kidder, E.N. Lewis, NIR News 18 (2007) 6–8.
[68] F.C. Clarke, S.V. Hammond, R.D. Jee, A.C. Moffat, Appl. Spectrosc. 56 (2002) 1475–1483.
[69] D.A. Burns, E.W. Cuirczak, Handbook of Near-Infrared Analysis, second edition, revised and expanded, Marcel Dekker, New York, 2001.
[70] D.L. Massart, B.G.M. Vandeginste, S.N. Deming, Y. Michotte, L. Kaufman, Chemometrics: A Textbook, Elsevier Science, Amsterdam, 1988.
[71] D. Bertrand, E. Dufour, La spectroscopie infrarouge et ses applications analytiques, 2nd edition, Tech & doc, Paris, 2006.
[72] R.P. Goehner, Anal. Chem. 50 (1978) 1223–1225.
[73] D. Bertrand, E. Vigneau, in: D. Bertrand, E. Dufour (Eds.), La spectroscopie infrarouge et ses applications analytiques, Tech & doc, Paris, 2006, pp. 427–447.
[74] R.J. Barnes, M.S. Dhanoa, S.J. Lister, Appl. Spectrosc. 43 (1989) 772–777.
[75] P. Geladi, D. MacDougall, H. Martens, Appl. Spectrosc. 39 (1985) 491–500.
[76] J. Burger, P. Geladi, J. Near Infrared Spectrosc. 15 (2007) 29–37.
[77] A. Kohler, C. Kirschner, A. Oust, H. Martens, Appl. Spectrosc. 59 (2005) 707–716.
[78] H. Martens, E. Stark, J. Pharm. Biomed. Anal. 9 (1991) 625–635.
[79] T.T. Cai, D. Zhang, D. Ben-Amotz, Appl. Spectrosc. 55 (2001) 1124–1130.
[80] Y.-P. Wang, Y. Wang, P. Spencer, Appl. Spectrosc. 60 (2006) 826–832.
[81] R. Bhargava, S.-Q. Wang, J.L. Koenig, Appl. Spectrosc. 54 (2000) 1690–1706.
[82] A.A. Green, M. Berman, P. Switzer, M.D. Craig, IEEE Trans. Geosci. Remote Sensing 26 (1988) 65–74.
[83] D. Bu, S.W. Huffman, J.A. Seelenbinder, C.W. Brown, Appl. Spectrosc. 59 (2005) 575–583.
[84] H. Takeuchi, S. Hashimoto, I. Harada, Appl. Spectrosc. 47 (1993) 129–131.
[85] C.J. Behrend, C.P. Tarnowski, M.D. Morris, Appl. Spectrosc. 56 (2002) 1458–1461.
[86] L. Zhang, M.J. Henson, Appl. Spectrosc. 61 (2007) 1015–1020.
[87] F. Vogt, S. Banerji, K. Booksh, J. Chemometr. 18 (2004) 350–362.
[88] F. Vogt, J. Cramer, K. Booksh, J. Chemometr. 19 (2005) 510–520.
[89] L.P. Choo, D.L. Wetzel, W.C. Halliday, M. Jackson, S.M. LeVine, H.H. Mantsch, Biophys. J. 71 (1996) 1672–1679.
[90] K.L.A. Chan, S.V. Hammond, S.G. Kazarian, Anal. Chem. 75 (2003) 2140–2146.
[91] K.L.A. Chan, S.G. Kazarian, J. Comb. Chem. 7 (2005) 185–189.
[92] K.L.A. Chan, S.G. Kazarian, Appl. Spectrosc. 61 (2007) 48–54.
[93] S. Šašic, D.A. Clark, J.C. Mitchell, M.J. Snowden, Analyst 129 (2004) 1001–1007.
[94] S. Šašic, Pharm. Res. 24 (2007) 58–65.
[95] H. Martens, T. Naes, Multivariate Calibration, John Wiley & Sons, Chichester, 1991.
[96] P. Paatero, U. Tapper, Chemometr. Intell. Lab. Syst. 18 (1993) 183–194.
[97] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classification, second edition, Wiley-Interscience, Chichester, 2004.
[98] S. Šašic, Appl. Spectrosc. 61 (2007) 239–250.
[99] S. Wold, M. Sjostrom, L. Eriksson, Chemometr. Intell. Lab. Syst. 58 (2001) 109–130.
[100] H. Martens, T. Naes, Multivariate Calibration, John Wiley & Sons, Chichester, 1991, 73 pp.
[101] J. Burger, P. Geladi, Analyst 131 (2006) 1152–1160.
[102] J. Burger, P. Geladi, J. Chemometr. 20 (2006) 106–119.
[103] L. Eriksson, E. Johansson, N. Kettaneh-Wold, S. Wold, Multi- and Megavariate Data analysis, Principles and Application, Umetrics Academy, Umea, 2001, 193 pp.
[104] N. Jovanovic, A. Gerich, A. Bouchard, W. Jiskoot, Pharm. Res. 23 (2006) 2002–2013.
[105] B.J. Westenberger, C.D. Ellison, A.S. Fussner, S. Jenney, R.E. Kolinski, T.G. Lipe, R.C. Lyon, T.W. Moore, L.K. Revelle, A.P. Smith, Int. J. Pharm. 306 (2005) 56–70.
[106] R. Tauler, A. Smilde, B. Kowalski, J. Chemometr. 9 (1995) 31–58.
[107] R. Tauler, Chemometr. Intell. Lab. Syst. 30 (1995) 133–146.
[108] W.H. Lawton, E.A. Sylvestre, Technometrics 13 (1971) 617–633.
[109] A. De Juan, R. Tauler, Anal. Chim. Acta 500 (2003) 195–210.
[110] J.-H. Jiang, Y. Liang, Y. Ozaki, Chemometr. Intell. Lab. Syst. 71 (2004) 1–12.
[111] E.R. Malinowski, Anal. Chim. Acta 134 (1982) 129–137.
[112] B.G.M. Vandeginste, W. Derks, G. Kateman, Anal. Chim. Acta 173 (1985) 253–264.
[113] E. Widjaja, M. Garland, J. Comput. Chem. 23 (2002) 911–919.
[114] E. Widjaja, R. Kim, H. Seah, J. Pharm. Biomed. Anal. 46 (2008) 274–281.
[115] W. Windig, J. Guilment, Anal. Chem. 63 (1991) 1425–1432.
[116] F.C. Sanchez, J. Toft, B. Van den Bogaert, D.L. Massart, Anal. Chem. 68 (1996) 79–85.
[117] A. Hyvärinen, J. Karhunen, E. Oja, Independant Component Analysis, first edition, John Wiley & Sons, Chichester, 2001.
[118] S. Moussaoui, C. Carteret, D. Brie, A. Mohammad-Djafari, Chemometr. Intell. Lab. Syst. 81 (2006) 137–148.
[119] C. Gobinet, E. Perrin, R. Huez, European Signal Processing Conference, Vienna, Austria, 2004.
[120] R. Bro, S. De Jong, J. Chemometr. 11 (1997) 393–401.
[121] D.D. Lee, H.S. Seung, Nature 401 (1999) 788–791.
[122] P. Sajda, S. Du, L. Parra, SPIE Wavelets x, San Diego, 2003.
[123] P. Paatero, Chemometr. Intell. Lab. Syst. 37 (1997) 23–35.
[124] P. Paatero, J. Comput. Graph. Stat. 8 (1999) 854–888.
[125] L. Duponchel, W. Elmi-Rayaleh, C. Ruckebusch, J.P. Huvenne, J. Chem. Inform. Comp. Sci. 43 (2003) 2057–2067.
[126] J.J. Andrew, T.M. Hancewicz, Appl. Spectrosc. 52 (1998) 797–807.
[127] C. Gendrin, Y. Roggo, C. Collet, J. Near Infrared Spectrosc. 16 (2008) 151–157.
[128] M. Maeder, Anal. Chem. 59 (1987) 527–530.
[129] A. De Juan, M. Maeder, T. Hancewicz, R. Tauler, Chemometr. Intell. Lab. Syst. 77 (2005) 64–74.
[130] A. De Juan, R. Tauler, Trends Anal. Chem.-TrAC 23 (2004) 70–79.
[131] J.-H. Wang, P.K. Hopke, T.M. Hancewicz, S.L. Zhang, Anal. Chim. Acta 476 (2003) 93–109.
[132] T. Hancewicz, J.-H. Wang, Chemometr. Intell. Lab. Syst. 77 (2005) 18–31.
[133] S.P. Gurden, E.M. Lage, C.G. De Faria, I. Joekes, M.M.C. Ferreira, J. Chemometr. 17 (2003) 400–412.
[134] J. Huang, H. Wium, K.B. Qvist, K.H. Esbensen, Chemometr. Intell. Lab. Syst. 66 (2003) 141–158.
[135] R. Bro, Chemometr. Intell. Lab. Syst. 38 (1997) 149–171.
[136] R. Bro, Multi-way Analysis in the Food Industry, University of Copenhagen, 1998.
[137] T. Isaksson, A.H. Aastveit, in: J.M. Chalmers, P.R. Griffiths (Eds.), Handbook of Vibrational Spectroscopy, vol. 3, John Wiley & Sons, London, 2002, pp. 2107–2122.
[138] A.K. Jain, M.N. Murty, P.J. Flynn, ACM Comput. Surv. 31 (1999) 264–323.
[139] A.K. Jain, R.P.W. Duin, J. Mao, IEEE Trans. Pattern Anal. Mach. Intell. 22 (2000) 4–37.
[140] C.T. Zugates, P.J. Treado, Int. J. Vib. Spectrosc. (1999).
[141] R. Bhargava, D.C. Fernandez, S.M. Hewitt, I.W. Levin, Biochim. Biophys. Acta (BBA)-Biomembr. 1758 (2006) 830–845.
[142] P. Geladi, H. Grahn, Multivariate Image Analysis, second edition, John Wiley & Sons, Chichester, 1996.
[143] M.H. Bharati, J.F. MacGregor, Ind. Eng. Chem. Res. 37 (1998) 4715–4724.
[144] H. Yu, J.F. MacGregor, G. Haarsma, W. Bourg, Ind. Eng. Chem. Res. 42 (2003) 3036–3044.

[145] W.H.A.M. van den Broek, D. Wienke, W.J. Melssen, L.M.C. Buydens, Anal. Chim. Acta 361 (1998) 161–176.
[146] P. Lasch, M. Diem, W. Hänsch, D. Naumann, J. Chemometr. 20 (2006) 209–220.
[147] J.A. Fernandez Pierna, V. Baeten, P. Dardenne, Chemometr. Intell. Lab. Syst. 84 (2006) 114–118.
[148] J.J. Liu, M.H. Bharati, K.G. Dunn, J.F. MacGregor, Chemometr. Intell. Lab. Syst. 79 (2005) 42–54.
[149] T.N. Tran, R. Wehrens, L.M.C. Buydens, Chemometr. Intell. Lab. Syst. 77 (2005) 3–17.
[150] J.R. Mansfield, M.G. Sowa, G.B. Scarth, R.L. Somorjai, H.H. Mantsch, Anal. Chem. 69 (1997) 3370–3374.
[151] J.R. Mansfield, M.G. Sowa, J.R. Payette, B. Abdulrauf, M.F. Stranc, H.H. Manisch, IEEE Trans. Med. Imag. 17 (1998) 1011–1018.
[152] J.C. Noordam, W.H.A.M. van den Broek, L.M.C. Buydens, Chemometr. Intell. Lab. Syst. 64 (2002) 65–78.
[153] P. Lasch, W. Haensch, D. Naumann, M. Diem, Biochim. Biophys. Acta (BBA)-Mol. Basis Dis. 1688 (2004) 176–186.
[154] S. Šašic, M. Morimoto, M. Otsuka, Y. Ozaki, Vib. Spectrosc. 37 (2005) 217–224.
[155] S. Šašic, D.A. Clark, J.C. Mitchell, M.J. Snowden, Appl. Spectrosc. 59 (2005) 630–638.
[156] J.F. Turner, J. Zhang, A. O'Connor, Appl. Spectrosc. 58 (2004) 1308–1317.
[157] P.W.T. Krooshof, G.J. Postma, W.J. Melssen, L.M.C. Buydens, T.N. Tran, Trends Anal. Chem. 25 (2006) 1067–1080.
[158] W.-Q. Lin, J.-H. Jiang, H.-F. Yang, Y. Ozaki, G.-L. Shen, R.-Q. Yu, Anal. Chem. 78 (2006) 6003–6011.
[159] J.C. Russ, The Image Processing Handbook, fourth edition, CRC Press, Boca Raton, 2002.
[160] E.N. Lewis, in: K.A. Bakeev (Ed.), Process Analytical Technology, Blackwell Publishing, Oxford, 2005, pp. 187–225.
[161] C. Gendrin, Y. Roggo, C. Collet, Eur. J. Pharm. Biopharm. 68 (2008) 376–385.
[162] N. Otsu, IEEE Trans. Syst. Man Cybern. 9 (1979) 62–66.
[163] W. Li, A. Woldu, R. Kelly, J. McCool, R. Bruce, H. Rasmussen, J. Cunningham, D. Winstead, Int. J. Pharm. 350 (2008) 369–373.
[164] F. Laplant, Am. Pharm. Rev. 7 (2004) 16–24.
[165] H. Ma, C.A. Anderson, J. Near Infrared Spectrosc. 15 (2007) 137–151.
[166] E.N. Lewis, J. Schoppelrei, E. Lee, Spectroscopy 19 (2004) 26–36.
[167] F.W. Koehler, e. Lee, l.H. Kidder, E.N. Lewis, Spectrosc. Eur. 14 (2002) 12–19.
[168] E.N. Lewis, J.E. Carrol, F.C. Clarke, NIR News 12 (2001) 16–18.
[169] S. Šašic, Anal. Chim. Acta 611 (2008) 73–79.
[170] L. Zhang, M.J. Henson, S.S. Sekulic, Anal. Chim. Acta 545 (2005) 262–278.
[171] A. El-Hagrasy, H.R. Morris, F. D'Amico, R.A. Lodder, J.K. Drennen, J. Pharm. Sci. 90 (2001) 1298–1307.
[172] R.C. Lyon, D.S. Lester, E.N. Lewis, E. Lee, L.X. Yu, E.H. Jefferson, A.S. Hussain, AAPS PharmSciTech 3 (2002) 1–17.
[173] E.N. Lewis, l.H. Kidder, E. Lee, Innov. Pharm. Technol. (2005) 107–111.
[174] http://www.in-pharmatechnologist.com/news/ng.asp?n=57633-chemical-imaging-investigated.
[175] O. Svensson, K. Abrahamsson, J. Engelbrektsson, M. Nicholas, H. Wikstrom, M. Josefson, Chemometr. Intell. Lab. Syst. 84 (2006) 3–8.
[176] C. Gendrin, Y. Roggo, C. Collet, Talanta 73 (2007) 733–741.
[177] E. Lee, W.X. Huang, P. Chen, E.N. Lewis, R.V. Vivilecchia, Spectroscopy (2006).
[178] A. De Juan, R. Tauler, Trends Anal. Chem. 23 (2004) 70–79.
[179] M.J. Henson, L. Zhang, Appl. Spectrosc. 60 (2006) 1247–1255.
[180] S. Ward, M. Perkins, J. Zhang, C.J. Roberts, C.E. Madden, S.Y. Luk, N. Patel, S.J. Ebbens, Pharm. Res. 22 (2005) 1195–1202.
[181] D.W. Rafferty, J.L. Koenig, J. Control. Rel. 83 (2002) 29–39.
[182] C.A. Coutts-Lendon, N.A. Wright, E.V. Mieso, J.L. Koenig, J. Control. Rel. 93 (2003) 223–248.
[183] S.G. Kazarian, K.L.A. Chan, Macromolecules 36 (2003) 9866–9872.
[184] J.v.d. Weerd, K.L.A. Chan, S.G. Kazarian, Vib. Spectrosc. 35 (2004) 9–13.
[185] J.v.d. Weerd, S.G. Kazarian, Appl. Spectrosc. 58 (2004) 1413–1419.
[186] J.v.d. Weerd, S.G. Kazarian, J. Control. Rel. 98 (2004) 295–305.
[187] J.v.d. Weerd, S.G. Kazarian, J. Pharm. Sci. 94 (2005) 2096–2109.
[188] N. Elkhider, K.L.A. Chan, S.C. Kazarian, J. Pharm. Sci. 96 (2007) 351–360.
[189] Y. Roggo, A. Edmond, P. Chalus, M. Ulmschneider, Anal. Chim. Acta 535 (2005) 79–87.
[190] L. Hilden, C.J. Pommier, S. Badawy, E.M. Friedman, Int. J. Pharm. 353 (2008) 283–290.
[191] R.B. Shah, M.A. Tawakkul, M.A. Khan, J. Pharm. Sci. 96 (2007) 1356–1365.
[192] C.D. Ellison, B.J. Ennis, M.L. Hamad, R.C. Lyon, J. Pharm. Biomed. Anal. 48 (2008) 1–7.
[193] W. Doub, W. Adams, J. Spencer, L. Buhse, M. Nelson, P. Treado, Pharm. Res. 24 (2007) 934–945.
[194] O.Y. Rodionova, L.P. Houmoller, A.L. Pomerantsev, P. Geladi, J. Burger, V.L. Dorofeyev, A.P. Arzamastsev, Anal. Chim. Acta 549 (2005) 151–158.
[195] C. Ricci, L. Nyadong, F. Fernandez, P. Newton, S. Kazarian, Anal. Bioanal. Chem. 387 (2007) 551–559.
[196] M.A. Veronin, E. Lee, E.N. Lewis, Ann. Pharmacother. 41 (2007) 1111–1115.
[197] M.A. Veronin, B.-B.C. Youan, Science 305 (2004) 481.
[198] I. Malik, M. Poonacha, J. Moses, R.A. Lodder, AAPS PharmSciTech 2 (2001) 1–7.
[199] S.J. Hamilton, A.E. Lowell, R.A. Lodder, J. Biomed. Opt. 7 (2002) 561–570.