Journal of Pharmaceutical and Biomedical Analysis 101 (2014) 123–140

Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

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

Review

# Data processing of vibrational chemical imaging for pharmaceutical applications

P.-Y. Sacré*, C. De Bleye, P.-F. Chavez, L. Netchacovitch, Ph. Hubert, E. Ziemons

*University of Liege (ULg), Department of Pharmacy, CIRM, Laboratory of Analytical Chemistry, CHU, B36, 4000 Liege, Belgium*

ARTICLE INFO

*Article history:*
Received 20 February 2014
Received in revised form 8 April 2014
Accepted 9 April 2014
Available online 19 April 2014

*Keywords:*
Hyperspectral imaging
Raman
Near-infrared
Infrared
Chemometrics

ABSTRACT

Vibrational spectroscopy (MIR, NIR and Raman) based hyperspectral imaging is one of the most powerful tools to analyze pharmaceutical preparation. Indeed, it combines the advantages of vibrational spectroscopy to imaging techniques and allows therefore the visualization of distribution of compounds or crystallization processes. However, these techniques provide a huge amount of data that must be processed to extract the relevant information.

This review presents fundamental concepts of hyperspectral imaging, the basic theory of the most used chemometric tools used to pre-process, process and post-process the generated data. The last part of the present paper focuses on pharmaceutical applications of hyperspectral imaging and highlights the data processing approaches to enable the reader making the best choice among the different tools available.



**Contents**



*Abbreviations:* ANN, artificial neural networks; API, active pharmaceutical ingredient; ATR-MIR, attenuated total reflectance mid-infrared; AUC, area under the curve; BTEM, band-target entropy minimization; CCD, charged coupled device; CLS, classical least squares; DHI, distributional homogeneity index; DPFT, dark point fixed transform; DR, depth resolution; EMSC, extended multiplicative scatter correction; FWHM, full width at half maximum; ICA, independent component analysis; k-NN, k-nearest neighbors; LDA, linear discriminant analysis; LIS, linear image signature; LIS-MVA, linear image signature multivariate data analysis; MCR-ALS, multivariate curve resolution-alternating least squares; MCT, mercury cadmium telluride; MIR, mid-infrared; MLR, multilinear regression; MSC, multiplicative scatter correction; MVA, multivariate data analysis; NIR, near-infrared; OSC, orthogonal signal correction; PARAFAC, parallel factor analysis; PAT, process analytical technology; PC, principal component; PCA, principal component analysis; PLS, partial least squares regression; PLS-DA, partial least squares discriminant analysis; PMF, positive matrix factorization; PSD, particle size distribution; RGB, red-green-blue; RMSD, root mean square difference; RMSE, root mean square error; ROI, region of interest; RPD, ratio performance deviation; RSD, relative standard deviation; SDBP, standard deviation between pixels; SECV, standard error of cross-validation; SEDDS, self-emulsifying drug delivery system; SG, Savitzky-Golay; SIMCA, soft independent modeling of class analogy; SIMPLISMA, simple-to-use interactive self-modeling mixture analysis; SISAL, simplex identification via split augmented Lagrangian; SMF, spectral match filter; SMMA, self-modeling mixture analysis; SNR, signal-to-noise ratio; SNV, standard normal variate; SVD, singular value decomposition; SVM, support vector machines; TTFA, target transformation factor analysis; XRPD, X-ray powder diffraction.

* Corresponding author at: Laboratory of Analytical Chemistry, CIRM, University of Liege (ULg), 4000 Liege, Belgium. Tel.: +32 43664324; fax: +32 43664317.
*E-mail address:* pysacre@ulg.ac.be (P.-Y. Sacré).

http://dx.doi.org/10.1016/j.jpba.2014.04.012
0731-7085/

## 1. Introduction

First developed in the environmental and remote sensing fields, chemical imaging also called hyperspectral imaging is now commonly used in a variety of sectors such as material, food, biological and pharmaceutical analysis [1].

The main interest of chemical imaging is that it combines both spatial and spectral information. It is therefore possible to observe things invisible to the naked eye such as distribution of components based on their chemical properties (different compounds, polymorphic forms, salts). The advantage of these properties appears clearly in the frame of pharmaceutical forms development and analysis. Indeed, most of the compounds used in the pharmaceutical industry are white or nearly white powders and it is impossible with naked eyes to observe the distribution of a white powder in a white powder.

Since the last decades, many initiatives have been set up in the pharmaceutical industries to get more understanding of processes with the objective of quality enhancement. To achieve this goal, a lot of research has been performed mainly with vibrational spectroscopic techniques such as Near-Infrared (NIR), Raman and Attenuated Total Reflectance-Mid-Infrared (ATR-MIR) [2–5]. Obviously, chemical imaging has its role to play within this frame and has been progressively more and more used for pharmaceutical dosage forms analysis [6].

Depending on the information sought, the appropriate spectroscopic technique and the appropriate data handling should be used. Indeed, each technique has its specificities, advantages and drawbacks that must be evaluated before starting any experiment. Once obtained, the collected data will encompass a specific pretreatment and analysis depending on the information that should be retrieved and interpreted.

The objective of this review is to present the most used chemical imaging techniques for pharmaceutical products analysis. Theoretical background of the most used chemometric techniques will be provided. Finally, a review of the literature concerning pharmaceutical applications of vibrational spectroscopy-based chemical imaging techniques will be presented with a focus on the specific data handling performed for each application or technique. We wrote this review to provide the reader basic guidelines to help him performing state of the art chemical imaging analysis of pharmaceutical products.

## 2. Fundamental concepts

### 2.1. Instrumentation for chemical imaging

NIR and Raman devices for chemical imaging are composed of a light source, optical parts (microscope objectives or lenses), a splitter (interferometer for Fourier-Transform devices, diffraction grating for dispersive devices or a tunable filter for global imaging devices) and a detector. MIR chemical imaging devices are mostly ATR devices with a germanium crystal and a mercury cadmium telluride (MCT) detector. All the spectroscopic parts of the device may be tuned for specific applications depending on the kind of sample studied and the sought information. More information about technical properties and tuning of devices may be found elsewhere [1,7,8].

#### 2.1.1. Configurations for chemical images acquisition

Chemical imaging devices may be arranged in three configurations: point mapping, line scanning and global imaging [6,9,10]. These configurations may exist with both macroscopic and microscopic imaging devices [11,12].

a. *Point mapping* (Fig. 1a). This configuration is probably the most used for Raman chemical imaging. It consists of recording a spectrum at a specific spatial location, then the sample is moved, another spectrum is recorded at location adjacent to the previous one and so on until the whole mapping area is covered. The main advantage of this technique is the high spectral quality (resolution, full spectral range) achievable since all parameters of spectral acquisition may be optimized. However, the main drawback is the acquisition time limited by the acquisition time and the moving of the sample. Nevertheless, new Raman spectrometers allow ultra-fast acquisition modes with acquisition times around 1 ms with acceptable signal-to-noise ratio (SNR).
b. *Line scanning* (Fig. 1b). Also known as the push-broom configuration, this technique is directly inspired from the remote sensing field. As with point mapping, spectra are recorded at defined spatial locations but with this configuration, the spectra of a whole line of the mapping area are simultaneously recorded. The advantages are the same as point mapping but the acquisition time is reduced.
c. *Global imaging* (Fig. 1c). This configuration is probably the most used in NIR imaging devices. The sample is entirely illuminated with a high power light source and the whole mapping area is recorded simultaneously at a single wavelength. Then the tunable filter moves, selects another wavelength and so on until the whole spectral range is recorded. The main advantage of this configuration is the fast acquisition time allowing the acquisition of a whole map in seconds. However, the spatial resolution is limited by the size (pixels) of the detector and the spectral resolution is limited by the type of tunable filters used. Furthermore, the heat produced by the high power light source may be a limiting factor with some kind of samples as it may melt them down or degrade the compound of interest.

As one can see, each configuration has its advantages and drawbacks. Unfortunately, these configurations are built-in and the analyst must choose a device that will meet its expectations based on a priori knowledge of the studied problem.

#### 2.1.2. Sampling and spatial resolution consideration

Spatial resolution may firstly be limited by the resolution power of the instrument (e.g. laser spot of ca. 100 μm for a macroscopic Raman system). But, when using microscopes, analysts want to have the highest possible spatial resolution. This resolution is, however, limited by the diffraction limit given by the equation:

$$r = 0.61\frac{\lambda}{NA} \quad (1)$$