IMAGING TECHNIQUES: A RAPID TOOL FOR FOOD ANALYSIS

Giorgia Foca
Department of Life Sciences and Interdepartmental Research Centre BIOGEST-SITEIA,
University of Modena and Reggio Emilia, Reggio Emilia, Italy, giorgia.foca@unimore.it

Abstract – In the last few decades there has been a continuous increase of industrial applications of image analysis-based techniques, thanks to their ability to monitor quickly and inexpensively both products and processes, and with the advantage of being non-invasive.
When performing image analysis, chemometrics is an essential tool for different aims, including data reduction or compression, extraction of useful information, and efficient representation of the most chemically-meaningful features of the analysed scene.

Keywords: imaging, RGB, hyperspectral, chemometrics, data reduction

1. BASICS ON IMAGING TECHNIQUES

The increasingly normative severity and market competitiveness have led food industry to constantly look for improvements of process monitoring systems. In the context of fast, non-destructive and reliable techniques, image analysis-based methods have gained particular interest, thanks to their ability to spatially characterise heterogeneous samples.

1.1. RGB imaging

Considering that food appearance strongly affects consumers' purchasing decisions, parameters such as size, shape, colour and presence/absence of visual defects are often used to monitor the quality throughout the production chain.

In this context, inspection by human operators has long been used for quality control and grading. Notwithstanding the proficiency of trained assessors, human assessment results to be quite subjective, laborious, expensive and not sufficient to meet the high throughput demands. Moreover, the human visual evaluation is operator dependent, and therefore not easily transferable between laboratories, industries or different production chains.

In order to minimise human intervention and to obtain an objective and reproducible sample evaluation, automated systems based on computer vision techniques have found wide applications in food industry [1]. Thanks to the spatial information which they are able to provide, image-based systems are in fact particularly suitable for the characterization of food matrices, where an inhomogeneous composition is often observed.

These systems are essentially based on the idea of reproducing and displaying a given scene in a way that simulates the human perception of colours. Indeed, the human imaging system, i.e., the eye, sees colour images by means of an array (the retina) of three kinds of sensors (the cone cells), which are sensitive to three different regions of the electromagnetic spectrum, perceived as the red, green and blue colours, respectively.
Similarly, artificial systems for colour image acquisition, such as photo- and video-cameras, flatbed scanners or webcams, are based on a sensor array (generally the charge-coupled device, CCD) which records in each picture element (pixel) the intensity values of red (R, at $\lambda \approx 630$ nm), green (G, at $\lambda \approx 545$ nm) and blue (B, at $\lambda \approx 435$ nm) spectral channels, in a manner to simulate the spectral response of the human eye. The resulting RGB image can be viewed as a three-dimensional data array, consisting of a given number of pixels (r rows × c columns), each one characterized by the three spectral variables Red, Green and Blue (Fig. 1).

Based on this approach, image processing systems able to determine external attributes such as size, shape, colour, surface texture and external defects of a wide range of food products have been developed over the years [2]. In addition, the possibility to use colour imaging to predict different food properties has also been widely investigated [3].

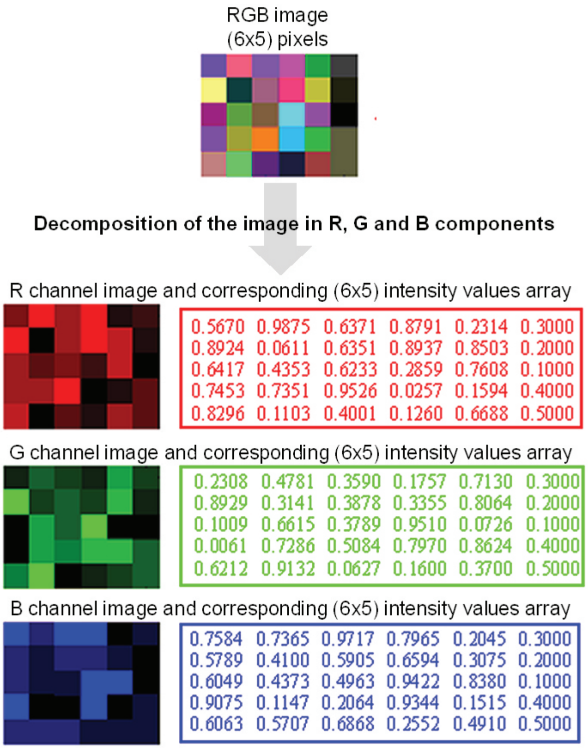

Fig. 1. Representation of the content of a RGB image.

1.2. Hyperspectral imaging

The simplest imaging systems are based on RGB images, however, for some applications, more than three variables and/or information not lying in the visible part of the electromagnetic spectrum can be necessary. In this context, an extension of RGB digital images which considerably widens the possibility to investigate the chemical-related aspects of the analyzed samples consists in hyperspectral imaging (HSI) systems, also known as chemical or spectroscopic imaging systems.

The use of HSI is rapidly emerging in the field of analytical chemistry, since it allows to collect simultaneously and in short times spectral and spatial information of a sample, and to represent it in the form of image-like maps of the chemical components distribution.

While RGB imaging systems record the intensity values of R, G and B channels for each pixel of the image, hyperspectral imaging systems collect the sample image over many spectral channels at different wavelengths (typically in visible and near infrared spectral ranges).

Compared with spectroscopic methods, HSI provides additional information by way of the simultaneous acquisition of spatially resolved spectra, combining the advantages of digital imaging with the attributes of spectroscopic measurements.

The data array resulting from HSI measurements is often referred to as *hypercube* (Fig. 2), given the three-dimensional nature of the multivariate data, with two spatial dimensions (*y* for pixel rows and *x* for pixel columns) and one spectral dimension (*λ* wavelengths) [4].

Therefore, HSI is not only useful to identify what chemical species are present in the sample and their concentration, but it also helps to highlight their location. This is a fundamental issue whenever the functionality of natural or artificial products depends on the spatial distribution of sample components, which is often the case in food, or when it is necessary to screen and identify contaminants in homogeneous samples.

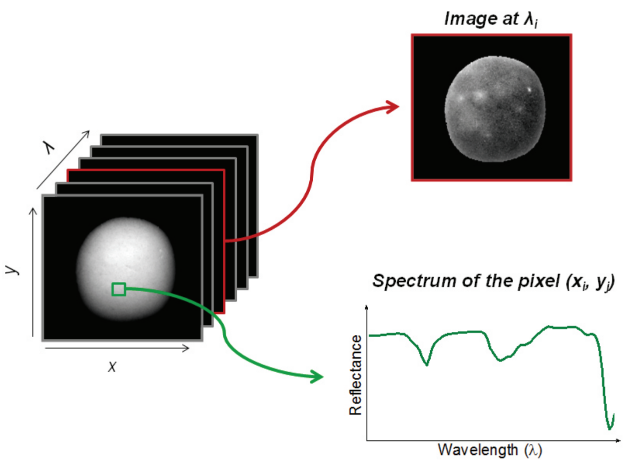

Fig. 2. Schematic representation of the structure of a hyperspectral image.

In spite of the major cost implied by hyperspectral systems in contrast with RGB ones, they generally provide much better results due to their greater potentials. For this reason, hyperspectral imaging applications have recently gained an increasing interest in food industry [5].

2. THE ISSUE OF DATA REDUCTION

Chemometrics can be defined in simple terms as *"the chemical discipline that uses mathematical and statistical methods for the obtention in the optimal way of relevant information on material systems"* [6]. With regard to imaging, chemometrics can be exploited to perform different tasks, such as image correction, segmentation and/or calculation of mathematical models which enable to determine properties of the samples directly from the acquired image.

Before imaging can be efficiently used for on-line or at-line applications, a crucial issue – in particular for hyperspectral images – has to be faced. Besides the high computational load and time required for model computation, a further consequence of the high amount of data conveyed by hyperspectral images is the difficulty to easily investigate and handle datasets composed by a large number of hypercubes. This probably represents the main limitation in food industry, where the final product quality is influenced by several factors (e.g., harvest period or animal feeding) and therefore there is the necessity to analyse many samples to account for the considerable inter-sample variability.

Nowadays, a number of multivariate-data reduction methods are available for the effective extraction of useful information from images.

2.1. From RGB images to colourgrams

In the context of automated methods aimed to extract the useful information contained in RGB images, our research group has developed an algorithm for the extraction and quantification of colour-related information [7], allowing to compress the whole colour-related information content of the image in a 1D signal, which can be used for data exploration, calibration, classification, product/process monitoring, etc.

This is essentially done by merging together quantities calculated on the analysed image, mainly consisting of frequency distribution curves of different colour-related parameters calculated for each pixel, including Red, Green, Blue, Intensity, ratios between Red, Green, Blue and Intensity, Hue, Saturation, Lightness, and data deriving from the Principal Component Analysis (PCA) of the RGB image. All the calculated parameters are then joined to form a signal, named *colourgram*, which codes the colour content of the image (Fig.3).

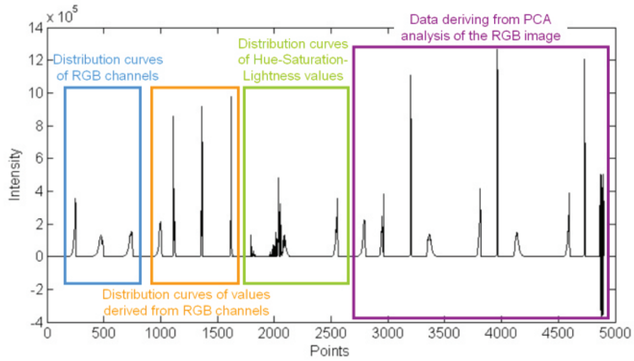

Fig. 3. A sample colourgram obtained from a RGB image, where the different groups of frequency distribution curves ("peaks") are highlighted.

This kind of data compression is particularly advantageous, since starting from the millions of data of the original images, elaboration is performed on a 4900 points long signal; colourgrams can then be considered as 'fingerprints' of the colour content of the samples. In this way, datasets made of multiple images can be analysed altogether: each single image is converted into the corresponding colourgram, and then proper multivariate analysis methods can be used to extract the information of interest from the whole dataset of colourgrams.

2.2. From hyperspectral images to hyperspectrograms

Dealing with HSI data, the most common and fast way applied to simultaneously compare all the hyperspectral images is to use the average (or median) spectra. This practice is very effective in the case of homogeneous samples, but do not perform properly when there is the need to identify spatially localised features within the image scene, since average spectra and median spectra do not account for spatial variability.

Following an approach similar to that used to derive the colourgrams starting from RGB images, our research group recently proposed a data reduction approach for hyperspectral images which is even compelling for heterogeneous samples [8]. Also in this case, the useful information contained in each hypercube is compressed into a 1D signal, named *hyperspectrogram*, which is derived from quantities extracted by means of PCA.

Essentially, the hyperspectrogram can be viewed as a fingerprint containing the relevant information brought by the hypercube, and is composed by a first part accounting for the spatial information and by a second part accounting for the spectral information. By representing each image with a vector of few hundreds of points, this procedure enables to simultaneously analyse up to hundreds of images by means of common multivariate analysis methods. Moreover, by applying proper variable selection procedures, hyperspectrograms can be further reduced to few significant descriptors, allowing to extract only the specific features that are useful to solve the problem at hand. Additionally, these features can be projected back into the image space, enabling a visual evaluation of the choices automatically made by the feature selection method, or into the spectral domain, in order to detect the spectral regions containing the information of interest (Fig. 4).

The use of colourgrams/hyperspectrograms to analyse datasets of images allows to: i) show similarities/dissimilarities between groups of images; ii) highlight the presence of defects or particular variations in some samples that could be difficult to recognise by visual inspection; iii) create calibration models, allowing to predict the value of specific properties of the sample, such as the content of a particular type of pigment, ingredient or contaminant; iii) create classification models, allowing to assign a sample to a class characterised by predetermined characteristics.

The effectiveness of this approach has been shown in the solution of different kinds of problems concerning different types of food samples. For instance, in a recent paper [9] we used colourgrams for the detection of the red skin defect in raw hams, in order to render more objective and transferable the evaluation usually made by expert assessors (Fig. 5).

An interesting application concerning hyperspectral data regards the early detection of bruises occurred on different apple varieties [10] (Fig. 6). Mechanical damages which can happen during fruit harvest, handling and transport represent a major cause of economic losses, and bruising in particular is the most common type of postharvest mechanical injury.

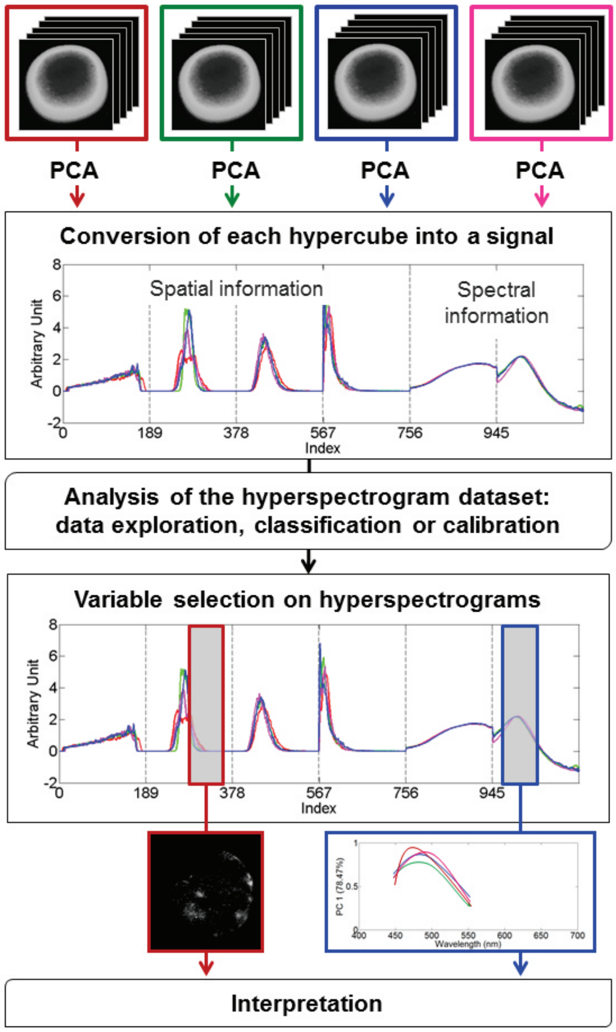

Fig. 4. Illustration of the hyperspectrogram approach.

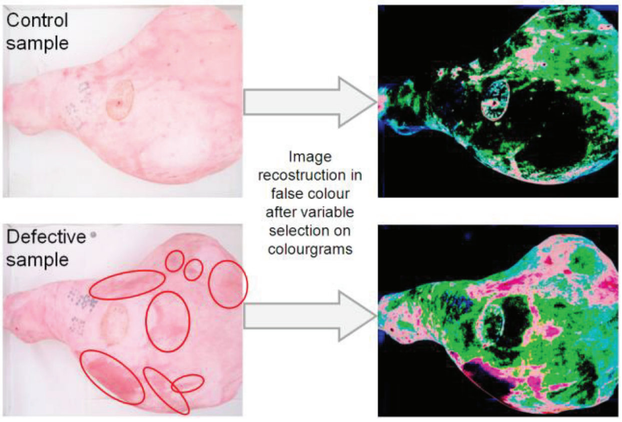

Fig. 5. Results of the application of colourgrams for the detection of red skin defect on hams.

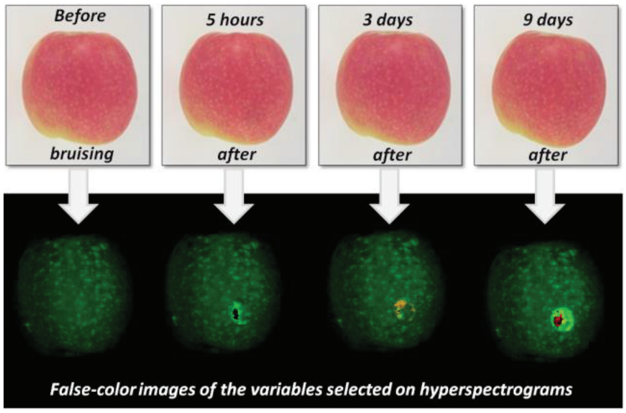

Fig. 6. RGB images and corresponding false-color images obtained on an apple before and after bruising.

3. CONCLUSIONS

Imaging is a novel platform technology that is been increasingly used in research field as well as in industry. It allows dealing with spatial features which enable the characterisation of complex heterogeneous samples and with spectral features which allow the identification of a wide range of surface constituents and sub-surface features. The development of on-line automated systems for image analysis is now particularly pursued in order to gain objective, accurate, rapid and non-destructive analysis. Future advances in imaging equipment manufacture, including lower purchase costs and enhancement in processing speed, will encourage more extensive utilisation of these emerging analytical methods.

REFERENCES

[1] D. Wu and D.-W. Sun, "Colour measurements by computer vision for food quality control – A review", *Trends in Food Science and Technology*, vol. 29, pp. 5-20, January 2013.

[2] H. Yu and J. F. MacGregor, "Multivariate image analysis and regression for prediction of coating content and distribution in the production of snack foods", *Chemometrics and Intelligent Laboratory Systems,* vol. 67, pp. 125-144, August 2003.

[3] V. Gökmen, H. Z. Şenyuva, B. Dülek and A. E. Çetin, "Computer vision-based image analysis for the estimation of acrylamide concentrations of potato chips and french fries", *Food Chemistry*, vol. 101, nº. 2, pp. 791-798, 2007.

[4] H. Grahn and P. Geladi, *Techniques and Applications of Hyperspectral Image Analysis,* John Wiley & Sons Ltd., Chichester, 2007.

[5] A. A. Gowen, C. P. O. Donnell, P .J. Cullen, G. Downey, J. M. Frias, "Hyperspectral imaging e an emerging process analytical tool for food quality and safety control", *Trends in Food Science and Technology,* vol. 18, pp. 590-598, 2007.

[6] I. E. Frank and B. R. Kowalski, "Chemometrics", *Analytical Chemistry*, vol. 54, pp. 232R-243R, 1982.

[7] A. Antonelli, M. Cocchi, P. Fava, G. Foca, G.C. Franchini, D. Manzini and A. Ulrici, "Automated evaluation of food colour by means of multivariate image analysis coupled to a wavelet-based classification algorithm", *Analytica Chimica Acta*, vol. 515, pp. 3-13, July 2004.

[8] C. Ferrari, G. Foca and A. Ulrici, "Handling large datasets of hyperspectral images : Reducing data size without loss of useful information", *Analytica Chimica Acta*, vol. 802, 29-39, November 2013.

[9] A. Ulrici, G. Foca, M.C. Ielo, L.A. Volpelli and D.P. Lo Fiego, "Automated identification and visualization of food defects using RGB imaging: Application to the detection of red skin defect of raw hams", *Innovative Food Science and Emerging Technologies*, vol. 16, 417–426, October 2012.

[10] C. Ferrari, G. Foca, R. Calvini and A. Ulrici, "Fast exploration and classification of large hyperspectral image datasets for early bruise detection on apples", *Chemometrics and Intelligent Laboratory Systems*, vol. 146, 108-119, May 2015.