Prediction of compositional and sensory characteristics using RGB digital images and multivariate calibration techniques / Foca, Giorgia; Masino, Francesca; Antonelli, Andrea; Ulrici, Alessandro. - In: ANALYTICA CHIMICA ACTA. - ISSN 0003-2670. - STAMPA. - 706:(2011), pp. 238-245. [10.1016/j.aca.2011.08.046]

(Article begins on next page)

PREDICTION OF COMPOSITIONAL AND SENSORY CHARACTERISTICS USING RGB DIGITAL IMAGES AND MULTIVARIATE CALIBRATION TECHNIQUES

G. Foca, F. Masino, A. Antonelli, A. Ulrici*

Dipartimento di Scienze Agrarie e degli Alimenti, Università degli Studi di Modena e Reggio Emilia, Padiglione Besta, via Amendola 2, 42122 Reggio Emilia, Italy.

* alessandro.ulrici@unimore.it
Telephone: +39 0522 522043
Fax: +39 0522 522027

Abstract

In the present paper, the possibility to use the information contained in RGB digital images to gain a fast and inexpensive quantification of colour-related properties of food is explored. To this aim, we present an approach which consists, as first step, in condensing the colour related information contained in RGB digital images of the analysed samples in one-dimensional signals, named *colourgrams*. These signals are then used as descriptor variables in multivariate calibration models. The feasibility of this approach has been tested using as a benchmark a series of samples of pesto sauce, whose RGB images have been used to predict both visual attributes defined by a panel test and the content of various pigments (chlorophylls *a* and *b*, pheophytins *a* and *b*, β-carotene and lutein).The possibility to predict correctly the values of some of the studied parameters suggests the feasibility of this approach for fast monitoring of the main aspect-related properties of a food matrix. The values of the squared correlation coefficient computed in prediction on a test set (R^2_{Pred}) for green and yellow hues were greater than 0.75, while R^2_{Pred} values greater than 0.85 were obtained for the prediction of total chlorophylls content and of chlorophylls/pheophytins ratio. The great flexibility of this blind analysis method for the quantitative evaluation of colour related features of matrices with an inhomogeneous aspect suggests that it is possible to implement automated, objective, and transferable systems for fast monitoring of raw materials, different stages of the manufacture and end products, not necessarily for the food industry only.

Keywords: RGB digital image analysis; multivariate calibration; wavelet transform; colour; food aspect; pigments; sensory evaluation;

1. Introduction

Visual aspect is one of the most important parameters for the assessment of food quality. For this reason, food industry is more and more interested in optimising not only taste and nutritional characteristics of a food product, but also its appearance, which is a complex combination of different characteristics including colour, texture, defects, etc. The identification of objective methods able to quantify visual aspect, or to codify some characteristics like amounts and distribution of the colour, is therefore fundamental. For example, correlations with sensory evaluation scores could help to implement automated and transferable systems for on-line process control or to check the aspect of the final product, even at the supermarket. Moreover, the definition of mathematical models linking colour-related aspects to chemical parameters (e.g. pH values or concentration of pigments) could help to explain the underlying mechanisms responsible for food appearance.

This research field is even more attractive because the instrumentation required to reach these goals is accessible at very low costs. In fact, many different types of digital cameras, webcams and scanners are able to convert the visual aspect of a food matrix in a series of data, i.e. the digital Red, Green and Blue (RGB) image. The key point is the definition of proper automated methods able to extract from RGB images the useful information and to employ it for calibration, classification, process monitoring, etc.

Scientific literature reports the use of digital images for the calibration of various food properties. The major part of the published research works describes applications for the prediction of technological and sensory aspects by means of various regression tools, including Multiple Linear Regression (MLR), Partial Least Squares (PLS) and Artificial neural Networks (ANN), in order to model the relationships between the information contained in the digital images and the investigated properties [1-8]. The image features that constitute the matrix of descriptors in the regression models are generally extracted by means of Multivariate Image Analysis (MIA) based techniques, which essentially consist in PCA of the unfolded image matrix. These techniques are able to convert images in objective and transferable information [7-9]. In some research works, the extracted image features are selected using different methods, such as Genetic Algorithms (GA) [2, 3] or Wavelet Transform (WT) [4].

Moisture [4, 10, 11], sensory and mechanical parameters [3, 12, 13] are some of the most frequently studied food properties using image analysis based calibration methods. Automated procedures to estimate the number of bacteria or the yeast mass grown on a proper support have been also reported [14, 15].

Only few research works concerning the image-based quantification of food components using multivariate calibration [7, 16] have been reported. The estimate of the chemical characteristics of food based on RGB images seems a more difficult task, since few works are available in the literature. Some Authors have modelled various properties (ranging from the lipids and carotenoids content in wheat to the percentage of potato chip surface covered by seasoning [17-20]) by means of univariate methods, obtaining satisfactory results. However, these univariate methods are specifically designed for the problem at hand and rarely can be transferred to a different application.

This short literature survey gives an idea of the wide possibilities offered by RGB imaging as a tool for fast and non-destructive characterisation of food matrices. However, the development of versatile methods, able to extract from the images a set of descriptors independently of the specific nature of the analysed matrix, and that can be used to build calibration models for a wide set of response variables, is not straightforward. In fact, many of the cited applications are customised on a specific food matrix for the prediction of a restricted group of variables. For example, the procedures that require image segmentation generally involve the customisation of the image analysis method. In fact, segmentation is often based on a problem-specific criterion, in order to isolate the sample from the background or to extract the informative portion of the image [3, 4, 11, 12, 16, 21, 22]. In other research works the customisation involves pretreatments like denoising, filtering, scaling and transformation of the colour space, which use is driven by the specific nature of the problem to be solved [4, 5, 7, 10, 11, 18, 23].

Following these considerations, in order to create a more versatile approach, Antonelli et al. [24] have developed an algorithm for the extraction of the overall colour-related information of the image, which is coded in the form of a signal named *colourgram*. The approach essentially consists in calculating the frequency distribution curves of a series of colour related parameters: red, green, blue, hue, saturation, lightness, and scores from the Principal Component Analysis (PCA) of the unfolded image data. The frequency distribution curves and the loadings and eigenvectors from PCA are then merged in sequence to give a 4900 points-long one-dimensional signal, the colourgram, which codes the colour content of the image.

A dataset composed by this kind of signals can then be used as an input to multivariate methods. For example, the most interesting features of a set of images can be investigated simply by calculating a PCA model on the corresponding matrix of colourgrams. Similarly, a matrix of colourgrams can be used as a set of descriptor variables to build classification or calibration models.

Compared to other approaches described in the literature, the main advantage offered by colourgrams lies in the flexibility of their possible applications. This is mainly due to the fact that colourgrams contain the whole information content of two different colour spaces (RGB and HSI) and of quantities derived from the RGB space. Therefore, there are no *a priori* assumptions on the features bringing the useful information for the specific problem under investigation. On the contrary, an image fingerprint that reflects all the complex colour-related features typical of a given food matrix is considered.

Moreover, the conversion of the image in a one-dimensional signal before processing allows higher data compression and easier computation. In fact, starting from the millions of data of the original image, elaboration is performed on a 4900 points-long signal, which can be further significantly shortened, up to few units, using proper signal compression/feature selection methods. Moreover, the methods available for processing monodimensional signals are more numerous, widespread and fast than the algorithms for image analysis.

The first application of the colourgrams approach [24] was made on an Italian pasta sauce, namely *Pesto alla genovese*. This food matrix was chosen as a benchmark, both since it contains particles of different size and colour, thus showing an inhomogeneous aspect, and for its colour instability, mainly due to the degradation of chlorophylls to pheophytins, which causes a change of colour from bright green to dull green-brownish [25]. The aim of this first work was to classify the digital images based on different pesto brands. The use of a feature selection algorithm based on the Wavelet Transform (WT) [26] for the identification of the most significant regions of the colourgrams allowed to reach a 100% classification efficiency in the prediction of an external test set.

These satisfactory results encouraged us to go a step further, evaluating the possibility to employ colourgrams for calibration purposes. Therefore, in the present work, we used colourgrams to predict both sensory and chemical properties, merging the results of previous studies on sensory [27] and chemical analysis [28] with the information contained in RGB images. The same food matrix, pesto sauce, was used as a benchmark for the same reasons described above.

Multivariate calibration models were built using both PLS and a feature selection/calibration algorithm based on WT [29], namely Wavelet Interface for Linear Modelling Analysis (WILMA) [30, 31]. The WILMA algorithm takes advantage of the multi-scale characteristics of the wavelet transform, which permits to consider both the shape of the signal (i.e., its frequency content) and its local aspects, such as peak positions and discontinuities. The variable selection implemented in WILMA allows to choose only those colour-related characteristics, which are the most relevant ones for a specific calibration task. The selected regions are generally contiguous, allowing an interpretation of the model based on visual inspection of these portions of the colourgram. In addition, this allowed the reconstruction of sample images displaying only the pixels with RGB values corresponding to the selected regions. In other terms, the feature selection made by the algorithm has been represented directly on the original images. Therefore, even though the colourgram contains only colour related information, regardless of spatial information (i.e., the specific location of each pixel is lost in the colourgram), it is still possible to have an "image-like" reconstruction of the selected features, which also allows their interpretation in spatial terms.

The results of the use of colourgrams for calibration purposes are very encouraging, showing a wide range of possible applications. A common RGB camera, together with an appropriate data processing as described in this paper, is able to catch enough information to make it suitable for quality control both in on-line and in off-line applications.

2. Experimental

2.1 Sampling and acquisition of digital images

In this study, twenty-four jars of pesto sauce from ten different producers have been considered. For nine producers two jars from single batches, indicated with letters from A to I have been analysed, while one producer supplied six jars from three different batches, indicated with letters from J to L.

After opening the jars, subsamples were used for image acquisition. Then, a part was kept and stored in a dark place at 4 ◦C to be analysed in the following days for the determination of pigment concentration [28]. Another part was kept to fill identical glass pots, identified by a numerical code, which were stored in a dark place at 4 ◦C as well, to be used for the panel test sessions in the following days [27]. During this time no appreciable degradation occurs, because the samples are stable under these storage conditions. In fact, pigments degradation occurs during pesto processing, when the product is acidified and pasteurised to improve microbiological stability.

For each one of the twenty-four original pesto jars, four aliquots of about 50 ml were separately collected and then spread on a flat (10×10) cm^2 surface to an approximately constant 5 mm thickness, for subsequent digital image acquisition. A dataset of 96 digital images of pesto was acquired as 24 bit RGB (16.8 millions of colours) with a 1280 × 960 spatial resolution, using a common Fujifilm Finepix S5000 digital camera, and then transferred to a personal computer as *jpeg* compressed image files (compression ratio 8:1). The choice to use low spatial resolution compressed images (average file size 487 KB) instead of higher resolution raw images (average file size 9 MB) was made on the basis of preliminary trials, where the score plots obtained by the PCA on colourgrams of some sample images acquired both in raw and in compressed modes at different resolutions were compared, showing similar patterns. Even if raw images could give better results preserving all the potentially useful information, the usability of the method (in terms of data storage and of computational power requirements) was considered more important.

The image acquisition system consisted in a white painted illumination chamber, equipped with 8 × 25W equally spaced tungsten lamps (Philips 25 W 240 V SES Argenta Lustre). The digital camera was placed on the aperture on the top of the chamber, 40 cm above the sample. A scene area of about 10 cm × 10 cm was covered, which corresponds to the sample surface area, taking care to avoid the presence of background pixels in the image.

In order to evaluate the effect of possible variations of the illumination conditions, preliminary tests were performed on the acquired images. However, the correction of the raw RGB values by means of a standard has not led to improvements of the results. Reasonably, this is due to the fact the light source variations within the relatively short time interval needed for image acquisition have not caused effects detectable by the CCD device of the RGB camera.

2.2 Sensory and chemical analyses

Sixteen judges had been specifically trained on the visual appearance of pesto samples, where the different attributes used to define aspect-related properties of pesto sauce were discussed with the help of some examples. Descriptive analysis was carried out evaluating the attributes listed in Table 1 on 10 cm long continuous line scales. In addition to six visual attributes, the personal preference of the judge, which is a hedonic attribute, was also included. Even if hedonic scales are usually reserved for consumer populations greater than 30, this hedonic attribute was included anyway, considering it as a sort of dummy variable. In fact, the variance due to subjectivity of the single assessors would have made PR difficult to estimate.

Each one of the twelve batches of pesto was submitted four times to the evaluation of each panellist during three separate sessions. In each session, twelve different samples and four replicates were evaluated following a Latin squares design. Finally, the scoring of each sample was calculated as the mean of the four evaluations. For a more detailed description of the sensory analysis, see reference [27].

On the same batches subjected to sensory evaluation, a double determination of the amount of the main pigments has been conducted. In particular, the following species were quantified: chlorophylls *a* and *b* (Chl *a* and Chl *b*), lutein (Lut), β-carotene (Car), pheophytins *a* and *b* (Pht *a* and Pht *b*). All pigments were determined by Reversed-Phase High-Performance Liquid Chromatography (RP-HPLC), except for Car, which was directly quantified on the purified extract by Vis spectrophotometry [28].

Beyond the concentration values of the single pigments, some derived compositional characteristics were also considered in the subsequent calibration models: total chlorophylls (ChlTOT, defined as Chl *a* + Chl *b*), total pheophytins (PhtTOT, defined as Pht *a* + Pht *b*) and chlorophylls/pheophytins ratio (Chl/Pht, defined as ChlTOT/PhtTOT). This latter quantity was considered since it expresses the extent of chlorophylls degradation.

2.3 Extraction and quantification of colour-related information

The colour related information of each digital image was used to build the corresponding colourgram. To this aim, the three way array corresponding to the RGB image, having size {960, 1280, 3} (where 960 is the number of pixel rows, 1280 the number of pixel columns, and 3 corresponds to the R, G and B colour channels) is unfolded to a {(960 × 1280), 3} bidimensional matrix containing all the pixels in rows and the R, G and B channels in columns. Then, this matrix is expanded by adding a series of columns, corresponding to parameters derived by R, G and B: Lightness (L), defined as the sum of the three channel values, the relative colours (rR, rG and rB), defined as the ratio between each channel and L, and the Hue, Saturation and Intensity values of the HSI colour space. Moreover, three PCA models are calculated: the first one on the unfolded RGB data matrix without any data pretreatment (raw), the second one after meancentering, and the last one after autoscaling. The nine score vectors (three for each PCA model) are also added as further columns to the data matrix. Then, for each one of the 19 columns of the resulting data matrix, the corresponding 256 points-long frequency distribution curve is calculated. The 19 frequency distribution curves are then joined in sequence to form a unique vector, and at the end the values of the three loading vectors (nine points) and of the eigenvalues of the three Principal Components (PCs) are added for all the three PCA models, to form a one-dimensional signal, the colourgram, of length equal to (256×19 + 36) = 4900 points, which describes the colour properties of the image. For a more detailed description of the algorithm used to build the colourgrams, the reader is referred to Antonelli et al. [24]. The matrix composed by the colourgrams (*colourgrams matrix*) was then used as predictor matrix for calibration purposes.

We are aware that the colour representation of the images by means of the whole colourgram is redundant; on the other hand, it must be emphasised that the 4900 points of a colourgram constitute anyway a small number, if compared to the (1280 × 960 × 3) = 3,686,400 data values of the original image array. Most important, in the same manner as the useful chemical information for a specific task is present only in specific portions of a NIR spectrum, also in this case the most relevant parts of the colourgrams for a given problem can be selected by means of proper signal processing and feature selection methods.

A simpler approach was also tested, using a *reduced matrix*, where for each digital image only 40 descriptor variables are calculated, corresponding to the mean, median, range and standard deviation values of parameters of the colourgram: R, G, B, L, rR, rG, rB, H, S, and I.

The same objects subdivision between training and test set, each one containing 48 objects, was done both for the *colourgrams matrix* and for the *reduced matrix*: for each batch all the replicates of the first jar were included in the training set and all the replicates of the second jar were included in the test set.

2.4 Multivariate Calibration and Feature Selection

As a starting point, PLS1 models have been applied to the colourgrams matrix and to the reduced matrix as predictors for each one of the 16 response variables. The combinations of different column-wise and row-wise preprocessing methods were considered. In particular, first order derivative (D1), second order derivative (D2) and standard normal variate (SNV) were applied alone and in combination with mean centering (MNCN). The best preprocessing and the optimal number of Latent Variables (LVs) were chosen on the basis of the results in contiguous blocks (12 groups) cross-validation. This cross-validation method was chosen in a way that the four replicates of each sample were contemporarily deleted during each cycle of cross-validation. As for the reduced matrix, autoscaling (AUTO) was used for data preprocessing, and also in this case the optimal number of LVs was chosen based on contiguous blocks (12 groups) cross-validation.

Since the information contained in colourgrams could be overwhelmed by uninformative variation, a feature selection technique was also used to compute the calibration models on the colourgrams matrix.

To this aim, we used the WILMA algorithm, which is based on the Fast Wavelet Transform (FWT) [29], a decomposition method in the WT domain. As a first step, i.e. at the first decomposition level, FWT splits the low and the high frequency contents of the signal (in this case, of each colourgram) into two orthogonal and complementary sub-spaces, called approximation and detail vector, respectively. To this aim, a couple of filters (the high-pass and low-pass wavelet filters) are used. Then, at the subsequent decomposition levels, it recursively splits into the approximation and detail vectors the approximation vector of the previous decomposition level, using the same wavelet filters. Therefore, at a given decomposition level L, the signal is represented by the approximation vector at level L and by all the detail vectors from level L to level 1. Each one of these vectors, which are defined in the same domain of the original signal, can be considered as a filtered version of the original signal, where only a restricted frequency (scale) range is kept. In other words, the set of variables (namely the wavelet coefficients) that are obtained at the various decomposition levels represent the contributions to the analysed signal at each position (in the original domain) and frequency (or scale) value. This double representation, which is called signal multiresolution, allows an efficient separation among all the signal features, which in turn permits the selection of only those aspects that are the most relevant to model a given response variable.

The wavelet coefficients, that constitute a set of independent variables derived from the colourgrams matrix, are then used in WILMA for the selection of the subset leading to the best predictive performance of the derived PLS/MLR regression models, evaluated by cross-validation.

Schematically, the WILMA algorithm works as follows:

- the matrix of signals (colourgrams) is decomposed by means of FWT using a particular wavelet (i.e. a particular couple of wavelet filters) until its maximum level of decomposition;
- for each decomposition level (including level 0, which corresponds to the original signal) the wavelet coefficients are ranked according to a criterion chosen by the user; in

the present work, based on our previous experience, we have considered the squared covariance of each coefficient with the response variable;

- for each decomposition level, the optimal number of wavelet coefficients is iteratively selected using either MLR or PLS with a proper cross-validation procedure; in the present work, since we wanted to keep all the images (thus, all the colourgrams) of pesto from a given jar in the same cross-validation groups, we have used contiguous blocks cross-validation with 12 groups;
- the wavelet coefficients belonging to the decomposition level that furnishes the best results in cross-validation are selected to build the final (optimal) calibration model;
- for interpretative purposes, only the selected wavelet coefficients are reconstructed into the original domain, allowing to point out the signal regions that contain the useful information.

For a detailed description of the WILMA algorithm, the reader is referred to more comprehensive references [30, 31].

Since it is not possible to know in advance the combination of the WILMA parameters that leads to the best results, it is appropriate to cycle over different possible combinations in order to find the best calibration model. For this reason, 5 wavelets belonging to the symlet family (sym4÷sym8), and both the MLR and the PLS regression techniques were used. For all the sixteen modelled response variables, the combinations resulting from the parameters listed above were tested, leading to 10 cycles of calculation (5 wavelets × 2 regression methods). The optimal MLR and PLS regression models were selected based on their cross-validation performance; finally, they were validated using the external test set samples.

For an easier comparison among the performances obtained for the different response variables with the different calibration methods, we have reported the results using the R^2 statistics, defined as:

$$R^2 = 1 - \frac{\sum(y - \hat{y})^2}{V_y(n-1)} \qquad (1)$$

where y are the experimentally measured values of the considered response variable, V_y is the corresponding variance, ŷ are the values calculated (for R^2_{Cal}) or predicted (in cross-validation for R^2_{CV} and on the test set for R^2_{Pred}) by the model, and n is the number of considered objects (i.e. the number of the training set objects for R^2_{Cal} and R^2_{CV}, and of the test set objects for R^2_{Pred}).

For all the four methods used for calibration (PLS on colourgrams matrix, PLS on reduced matrix, WILMA-PLS and WILMA-MLR), the number of latent variables (PLS-based methods) or of selected variables (WILMA-MLR) that was included in the models was chosen based on the minimum error in cross-validation, up to a maximum value of 6.

3. Results and Discussion

Table 2 reports the performances of the best calibration models selected for each response variable with each calibration method. In particular, Table 2.*a* reports the PLS1 models obtained using as X block both the colourgrams matrix (on the left side), composed by 4900 variables, and the reduced matrix (on the right side), composed by 40 variables, while Table 2.*b* reports the results obtained with WILMA, considering separately the models selected with WILMA-PLS (on the left side) and with WILMA-MLR (on the right side). For a general comparison among the four calibration methods and to point out the response variables characterised by acceptable correlations with the information from digital images, the variable names and the model statistics for the method with best overall performance in prediction of the test set are highlighted in grey colour, when R^2_{Pred} is equal or greater than 0.6.

On the whole, acceptable models were obtained for 10 variables out of 16; better results were obtained for the chemical variables (pigments concentrations) rather than for the sensory ones.

As for the sensory attributes, the three colourgrams-based methods converged, giving good calibration models only for yellow and green hues (YH and GH), while the reduced matrix-based approach gave acceptable results also for white amount (WA). As for GH, these results are coherent with the observations made in a previous research work [27], where it was observed that this parameter has a prominent role in the definition of the multivariate structure of the sensory data. In other words, the contribution of GH to the overall aspect is much evident, therefore it is easier to quantify by means of digital imaging techniques.

Conversely, CH and PS cannot be predicted, regardless of the method used for calibration. To explain the failure of our approach in the prediction of these parameters, it must be recalled that colourgrams retain the colour information but lose the spatial information. This is the reason why is not possible to determine if the pixels of similar colour are grouped together in relatively large clusters (corresponding to coarse PS or low CH) or more uniformly distributed along the spatial dimensions (corresponding to fine PS or high

CH). The only texture-related information that could have been partly extrapolated by the colourgrams approach (and that justified our attempt to predict PS and CH) derives from "boundary" pixels, i.e. from those pixels corresponding to the boundaries between particles having different colour. Their "intermediate" colour properties could have allowed their distinction from the "bulk" pixels and therefore their use to estimate PS and CH. However, the experimental results confirm that this hypothesis is not verified, and that in order to obtain textural information with a colourgrams-based approach further work is needed, to insert also spatial information in the signals.

As expected, PR generally shows the worst performance in prediction (lowest R^2_{Pred} values). In fact, we recall that this hedonic attribute was considered as a sort of dummy variable, since the variance due to subjectivity of the single assessors would have made difficult to estimate PR based on images. Thus, these bad results in prediction can be considered as a sort of check of the correctness of the validation procedure.

As for the chemical variables, only those related to the pigments whose contribution to the overall colour (green) is more evident, i.e. pheophytins and chlorophylls, gave acceptable results. In particular, independently of the specific calibration method, the best models were obtained for the Chl/Pht ratio (highest R^2_{Pred} = 0.872 with WILMA-PLS) and for the total amount of chlorophylls, ChlTOT (highest R^2_{Pred} = 0.876 with WILMA-MLR).

Satisfactory models with all the calibration methods were obtained for Chl *a* and Chl *b*, with the only exception of Chl *a* modelled with the reduced matrix. Pht *a* and Pht *b* gave worse results, even though generally acceptable. Conversely, none of the four calibration methods gave satisfactory results for the quantification of Lut and of Car. On the one hand, this fact suggests that RGB images are not able to detect all the pigments, but only the ones with the higher contribution to the overall colour of the samples, at least under our experimental conditions. On the other hand, all the four calibration methods show a good degree of convergence in predicting the same chemical variables, suggesting that the approaches adopted for calibration are not prone to overfitting.

The performance of the different methods used to build the calibration models showed that the simplest approach, i.e. PLS on the reduced matrix, gave better results than expected. In fact, it led to the best models for three out of the sixteen response variables (WA, Pht *a* and PhtTOT). In particular, it is the only one able to correctly predict the values of WA. This fact suggests that the use of global parameters, like those of the reduced matrix, can be sometimes more effective than a detailed analysis of the frequency distribution curves (as it is done using WILMA).Figure 1 shows the VIP (Variable Importance in Projection) scores plot of the PLS regression model of WA based on the reduced matrix. The VIP scores estimate the importance of each variable in the projection used in a PLS model [32]; the "greater than one rule", which derives from the fact that the average of squared VIP scores equals 1, helps to determine whether a certain variable is actually significant to the model. From this figure it is clear that the useful information to estimate WA consists mainly in the mean and median values of the intensities of the single channels (R, G, and B, where G is the most important one, as it could be expected based on the green dominant colour of pesto), of their sum (as expressed by L), and of their maximum values (as expressed by I).

The global number of acceptable results ($R^2_{Pred} > 0.6$) obtained with the reduced matrix and with the colourgrams matrix is analogous. However, the overall performance in prediction of colourgrams-based models is better. In fact, with the reduced matrix only three variables lead to $R^2_{Pred} > 0.7$ with a maximum of $R^2_{Pred} = 0.773$ for ChlTOT, while using colourgrams the number of models with $R^2_{Pred} > 0.7$ ranges from 5 (PLS and WILMA-PLS) to 6 (WILMA-MLR) and the number of models with $R^2_{Pred} > 0.8$ ranges from 1 (PLS) to 3 (WILMA-PLS and WILMA-MLR).

As for the colourgrams-based models obtained with PLS, it is evident (Table 2.*a*) that there is not a specific pretreatment prevailing on the other ones, even though in general derivatives, alone or combined with mean centering, perform better than SNV.

Figure 2 reports the VIP scores plot of the PLS model on colourgrams for GH (Figure 2.*b*) together with a mean colourgram of the training set (Figure 2.*a*). Six distinct colourgram portions, corresponding to the frequency distribution curves of different parameters related to the colour content of the image, have a strong influence on the model (VIP score values much higher than 1). These portions, highlighted in grey colour, correspond to rG (1), H (2), PC3 scores of the PCA on raw (3), meancentered (4) and autoscaled (5) matrices, and to the loadings and eigenvalues for all the three PCA models (6) [24]. The selection of the frequency distribution curves of rG and H can be easily connected with the nature of the investigated property. The interpretation of the selected distribution curves of the PCA scores of the third component of the three PCA models and of the corresponding loadings and eigenvalues is more difficult, but it highlights the need to account for the inner relations between the R, G and B channels, in order to model properly the GH sensory attribute.

In general, the fact that only few narrow regions are necessary to model the response variable, suggested us how the implementation of feature selection techniques can be of great help in building efficient and at the same time parsimonious models, where – starting from the comprehensive description of the image colour content furnished by the colourgram – at the end only few descriptors can be extracted for a specific task. In fact, the results obtained by applying WILMA to the colourgrams matrix demonstrate that the selection of variables is useful, not only since this often led to models more efficient than those obtained with PLS, but also since the number of selected wavelet coefficients ("cfs" columns in Table 2.*b*) is frequently very low. Beyond the aspects connected to predictive performance and robustness, the advantages offered by the high reduction of useful descriptors is twofold. On the one hand, simple models based on few variables are computationally very fast, which makes easier their online implementation. On the other hand, the representation in the original images of only those pixels having values included in the intervals selected on colourgrams enables i) a direct visualisation of the choices made by the algorithm for model interpretation purposes and ii) the monitoring of each analysed image, e.g. to check for possible outlying samples or to check if an image defect affects the descriptors used for calibration.

All the five tested wavelets have been selected in the various models ("Wav" columns in Table 2.*b*), confirming that the optimal wavelet depends on the specific calibration task, even if sym6 and sym7 have been selected more frequently than the other ones. Also the optimal wavelet decomposition levels vary considerably ("Lev" columns in Table 2.*b*), including the extreme possibilities of the original signal (0) and of the maximum decomposition level (12).

As for the performance in prediction, the best results were obtained with WILMA-PLS, which gives acceptable results ($R^2_{Pred} > 0.6$) for nine out of the sixteen modelled variables. However, satisfactory results were obtained also with WILMA-MLR, with the advantage that the number of selected wavelet coefficients (descriptors) is extremely reduced, ranging from one to six. It is noteworthy the fact that, though based on different strategies for feature selection, WILMA-MLR and WILMA-PLS converge for both the PS and PR models, where the same unique wavelet coefficient is selected for each one of these response variables. The tendency to converge towards the same (bad) solution, without finding any possible chance correlation with a higher number of coefficients, demonstrates that the adopted feature selection / cross-validation procedure is not prone to overfitting.

The overall best results in prediction were obtained using WILMA-MLR for the model of total chlorophylls content (ChlTOT). Figure 3 reports both the original colourgrams (Figure 3.*a*) and the reconstruction of the selected wavelet coefficients into the original domain (Figure 3.*b*). The three portions of the colourgram that have been selected correspond to the frequency distribution curves of H (1) and of the PC3 scores of the PCA models on the meancentered (2) and on the autoscaled (3) matrices. The comparison between this Figure and Figure 2.b, that reports the VIP scores of the PLS model for GH, shows the success of the blind analysis approach, suggesting how different algorithms converge to analogous signal regions for correlated properties (the correlation coefficient between GH and ChlTOT = 0.8273).

The first one of the three regions highlighted in Figure 3 (H) is represented in more detail in Figure 4, where the training set objects are represented with black lines and the test set objects with grey lines. In Figure 4.*a* different shapes of the frequency distribution curves can be observed, which correspond to the different jars of pesto sauce. The reconstruction of the (unique) wavelet coefficient selected in the H region of the colourgram shows that this coefficient is located in a way to account for all the main differences between the various H frequency distribution curves.

Given the importance of this colourgram region to model ChlTOT, considered that it corresponds to a limited number of H values, we have then verified whether these values effectively correspond to image pixels pertaining to the parts of pesto sauce where chlorophylls are still present. To this aim, we have considered two images, one taken on a sample with a high ChlTOT value (sample K) and another one taken on a sample with a low ChlTOT value (sample G). For both images the H values were calculated for each pixel, and only those pixels having H values lying within the selected range were maintained, the remainder ones being set to black (R=G=B=0). Figure 5 reports the original images (on the left side) and the reconstructed ones (on the right side) for sample K (a1 and a2) and for sample G (b1 and b2). It is evident the high difference in the number of selected pixels for the two samples, which is the quantity that the calibration model correlates to ChlTOT. A more accurate comparison between the original and the reconstructed images reveals that the selected pixels essentially correspond to the green particles of the sample. This kind of representation, in a more explicit manner than Figure 4.b, allows to see what the algorithm actually considered (as for the contribution of H), and to “resume” the spatial information that in the colourgram was lost.

In view of possible implementations of the method for quality control in a laboratory and, after a proper engineering, also on a production line, images like those in Figure 5 give additional information for the interpretation of results. In addition, these outputs are also accessible to people not necessarily expert in the technical (chemometric) aspects of the image elaboration. Moreover, the variables selection process allows a drastic lowering of the number of informative features extracted by means of the colourgram. This fact suggests that, once the proper model able to work on a few variables has been developed, it should be possible to process each image in a very short time and contemporarily to monitor the analysed samples in real time.

4. Conclusions

The information relative to the colour of a sample, codified in the form of a signal extracted from the RGB image data, allowed to build reliable multivariate calibration models with several chemical and sensory properties of pesto sauce, which has been chosen as a benchmark food matrix. Although in this preliminary research work calibration models were obtained on a relatively small set of samples, satisfactory results were gained for the properties that mainly contribute to the sample aspect, i.e. the sensory properties GH and YH, and the chlorophylls and chlorophylls/pheophytins ratio as for the pigments content.

These results suggest the possibility to implement quantitative models in automated monitoring systems of raw materials, different stages of the manufacture and end products, to check the sensory quality and the state of preservation, which is reflected into the pigment amount.

It is important to stress that the application to the colourgrams of algorithms that perform feature selection allows to extract, from a signal containing a wide range of potentially useful information, only a small number of useful variables, without specific *a priori* assumptions on the types of features to be considered. This implies a great flexibility in the possible uses, including all applications on any kind of sample having an inhomogeneous aspect, where a colour-related problem has to be handled.

Considering the continuous improvements of the image data quality and of the commercially available computational power and data storage capabilities, the possibility to use raw images at higher resolutions could help in enhancing the quality of the analysed data and therefore of the calibration models.

Finally, the reported results also suggest possible further developments of this approach, like the inclusion into the colourgrams of additional information, which could include both global parameters such as means and/or standard deviations, and also other variables describing texture-related aspects of the image.

References

[1] M.M. Lana, L.M.M. Tijskens, O. van Kooten, J. Food Eng. 77 (2006) 871–879.

[2] P. Jackman, D.W. Sun, P. Allen, J. Food Eng. 96 (2010) 151–165.

[3] P. Jackman, D.W. Sun, C.J. Du, P. Allen, Pattern Recogn. 42 (2009) 751–763.

[4] M. Shafafi Zenoozian, S. Davahastin, J. Food Eng. 90 (2009) 219–227.

[5] S.M.B. Johansen, J.L. Laugesen, T. Janhøj, R.H. Ipsen, M.B. Frøst, Food Qual. Prefer. 19 (2008) 232–246.

[6] C.J. Du, D.W. Sun, P. Jackman, P. Allen, Meat Sci. 80 (2008) 1231–1237.

[7] H. Yu, J.F. MacGregor, Chemom. Intell. Lab. Syst. 67 (2003) 125–144.

[8] S. Garcia-Munoz, D.S. Gierer, Int. J. Pharm.395 (2010) 104–110.

[9] J.M. Prats-Montalban, A. Ferrer, R. Bro, T. Hancewicz, Chemom. Intell. Lab. Syst. 96 (2009) 6–13.

[10] M. Mohebbi, M.R. Akberzadeh-T, F. Shahidi, M. Moussavi, H.B. Ghoddusi, Comp. Electr. Agric. 69 (2009) 128–134.

[11] C. Zheng, D.W. Sun, L. Zheng, J. Food Eng. 77 (2006) 858–863.

[12] M.R. Chadraratne, S. Samarasinghe, D. Kulasiri, R. Bikerstaffe, J. Food Eng. 77 (2006) 492–499.

[13] J. Qiao, N. Wang, M.O. Ngadi, S.Kazemi, J. Food Eng. 79 (2007) 1065–1070.

[14] A. Borin, M. Flôres Ferrão, C. Mello, L. Cordi, L.C.M. Pataca, N. Durán, R.J. Poppi, Anal. Bional. Chem. 387 (2007) 1105–1112.

[15] C.A. Acevedo, O. Skurtys, M.E. Young, J. Enrione, F. Pedreschi, F. Osorio, LWT-Food Sci. Technol. 42 (2009) 1444–1449.

[16] P. Rajbhandari, P.S. Kindstedt, J. Dairy Sci. 88 (2005) 4157–4164.

[17] I. Konopka, W. Kozirok, D. Rotkiewicz, Food Res. Int. 37 (2004) 429–438.

[18] R.A. Quevedo, J.M. Aguilera, F. Pedreschi, Food Bioprocess. Technol. 3 (2010) 637–643.

[19] T. Zhou, A.D. Harrison, R. McKellar, J.C. Young, J. Odumeru, P. Piyasena, X. Lu, D.G. Mercer, S. Karr, Food Res. Int. 37 (2004) 875–881.

[20] C. Reyes, S.A. Barringer, J. Food Proc. Preserv. 29 (2005) 369–377.

[21] V. Gökmen, H.Z. Şenyuva, B. Dülek, A.E. Cetin, Food Chem. 101 (2007) 791–798.

[22] A. Mateo, F. Soto, J.A: Villarejo, J. Roca-Dorda, F. De la Gandara, A. García, Aquac. Eng. 35 (2006) 1–13.

[23] F. Pedreschi, J. León, D. Mery, P. Moyano, Food Res. Int. 39 (2006) 1092–1098.

[24] A. Antonelli, M. Cocchi, P. Fava, G. Foca, G.C. Franchini, D. Manzini, A. Ulrici, Anal. Chim. Acta 515 (2004) 3–13.

[25] T.P. Coultate, Colours, in: Food – The Chemistry of its Components, fourth ed., RCS Paperbacks, Cambridge, UK, 2002, pp. 175–217.

[26] M. Cocchi, R. Seeber, A. Ulrici, Chemom. Intell. Lab. Syst. 57 (2001) 97–119.

[27] F. Masino, G. Foca, A. Ulrici, L. Arru, A. Antonelli, J. Sci. Food Agr. 88 (8) (2008) 1335–1343.

[28] F. Masino, A. Ulrici, A. Antonelli, Eur. Food Res. Technol. 226 (2008) 569–575.

[29] B. Walczak (Ed.), Wavelets in Chemistry. Elsevier, Amsterdam, 2000.

[30] M. Cocchi, R. Seeber, A. Ulrici, J. Chemometrics 17 (2003) 512–517.

[31] A. Ulrici, M. Cocchi, G. Foca, C. Durante, A. Marchetti, L. Tassi, in: M.P. Colombini, L. Tassi (Eds.), New trends in analytical, environmental and cultural heritage chemistry, Research Signpost, Trivandrum, India, 2008, pp. 77–136.

[32] I.G. Chong, C.H. Jun, Chemom. Intell. Lab. Syst. 78 (2005) 103–112.

Captions to Tables and Figures

Table 1. List of the visual attributes, together with the corresponding abbreviations and ranges.

Table 2. Best calibration results *a*) from PLS models obtained on Colourgrams matrix and on reduced matrix, and *b*) from WILMA models obtained using PLS or MLR as regression methods on Colourgrams matrix. Grey background indicates the overall best model for a specific property. A separation between sensory and chemical variables is highlighted.

Figure 1. Plot of the VIP scores of the PLS model on the reduced matrix for the prediction of WA.

Figure 2. Green hue (PLS model): a) mean colourgram for training set; b) VIP scores where the regions that are significant to model GH are highlighted.

Figure 3. Selected signal regions for ChlTOT (WILMA-MLR model): a) original colourgrams; b) reconstructed signal where the regions significant to model ChlTOT are highlighted.

Figure 4. Zoom on the selected region n. 1 of Figure 3, corresponding to the frequency distribution curve of the Hue values. The rectangles plotted with dotted lines define the property intervals that were selected for the further image reconstruction.

Figure 5. Examples on K and G samples of image reconstruction for ChlTOT (WILMA-MLR model). Original sample images (a1 and b1), and corresponding images reconstructed with selected Hue values (a2 and b2).

TABLES

Abbreviation	**Variable name**	**Range (from *0* to *10*)**
GH	Green Hue	from *dull* to *bright*
YH	Yellow Hue	from *low* to *high*
BH	Brown Hue	from *low* to *high*
WA	White Amount	from *low* to *high*
CH	Colour Homogeneity	from *low* to *high*
PS	Particle Size	from *fine* to *coarse*
PR	PReference	from *low* to *high*

Table 1

a)

Variable	PLS on colourgrams matrix					PLS on reduced matrix				
	Pretreatment	LVs	R^2_{Cal}	R^2_{CV}	R^2_{Pred}	Pretreatment	LVs	R^2_{Cal}	R^2_{CV}	R^2_{Pred}
GH	D2	2	0.837	0.630	0.781	AUTO	6	0.967	0.843	0.674
YH	SNV	1	0.789	0.522	0.757	AUTO	6	0.976	0.750	0.696
BH	D1+MNCN	3	0.731	0.068	0.527	AUTO	2	0.595	0.281	0.366
WA	D1+MNCN	4	0.909	0.537	0.491	AUTO	3	0.837	0.632	0.683
CH	D1+MNCN	3	0.793	0.064	0.344	AUTO	5	0.853	0.333	0.402
PS	D2	1	0.282	-0.439	0.145	AUTO	6	0.895	-0.289	-0.413
PR	D1	1	0.282	-0.383	0.160	AUTO	6	0.877	-0.145	-1.945
Car	D2+MNCN	1	0.399	0.045	0.467	AUTO	6	0.942	0.481	0.509
Lut	SNV+MNCN	1	0.432	0.020	0.496	AUTO	5	0.908	0.339	0.516
Chl *a*	D2	2	0.748	0.495	0.650	AUTO	4	0.909	0.290	0.374
Chl *b*	D1	3	0.916	0.842	0.718	AUTO	1	0.845	0.725	0.732
ChlTOT	D2	6	0.979	0.877	0.733	AUTO	1	0.803	0.567	0.773
Pht *a*	D2	6	0.879	0.352	0.378	AUTO	5	0.944	0.667	0.665
Pht *b*	SNV+MNCN	1	0.567	0.273	0.607	AUTO	3	0.786	0.419	0.313
PhtTOT	D1	3	0.662	0.196	0.432	AUTO	5	0.925	0.574	0.641
Chl/Pht	D1	5	0.961	0.817	0.803	AUTO	4	0.972	0.822	0.738

b)

Variable	WILMA-PLS							WILMA-MLR					
	cfs	Wav	Lev	LVs	R^2_{Cal}	R^2_{CV}	R^2_{Pred}	cfs	Wav	Lev	R^2_{Cal}	R^2_{CV}	R^2_{Pred}
GH	34	sym6	3	1	0.865	0.765	0.745	4	sym8	1	0.888	0.817	0.777
YH	117	sym4	3	1	0.822	0.750	0.731	4	sym8	5	0.876	0.720	0.766
BH	43	sym7	5	3	0.741	0.325	0.554	6	sym4	2	0.599	0.113	-0.391
WA	20	sym4	9	3	0.901	0.698	0.469	4	sym7	3	0.840	0.697	-0.520
CH	1	sym7	1	1	0.329	0.078	0.298	3	sym8	5	0.739	0.428	-0.082
PS	1	sym6	2	1	0.258	0.047	0.239	1	sym6	2	0.258	0.047	0.239
PR	1	sym7	2	1	0.232	0.011	0.129	1	sym7	2	0.232	0.011	0.129
Car	767	sym7	3	2	0.790	0.283	0.206	4	sym6	1	0.625	0.327	0.322
Lut	767	sym7	3	2	0.787	0.306	0.286	2	sym4	0	0.436	0.229	0.509
Chl *a*	3	sym5	1	3	0.880	0.678	0.674	3	sym6	1	0.881	0.623	0.713
Chl *b*	24	sym5	3	1	0.926	0.892	0.803	1	sym4	2	0.901	0.828	0.802
ChlTOT	26	sym6	3	6	0.991	0.924	0.839	4	sym6	1	0.970	0.938	0.876
Pht *a*	122	sym6	2	1	0.615	0.506	0.627	4	sym7	1	0.702	0.506	0.495
Pht *b*	89	sym5	2	1	0.658	0.470	0.627	2	sym5	1	0.579	0.425	0.617
PhtTOT	150	sym8	2	1	0.602	0.498	0.608	4	sym4	0	0.685	0.478	0.542
Chl/Pht	13	sym7	12	5	0.972	0.919	0.872	3	sym6	1	0.978	0.935	0.872

Table 2

FIGURES

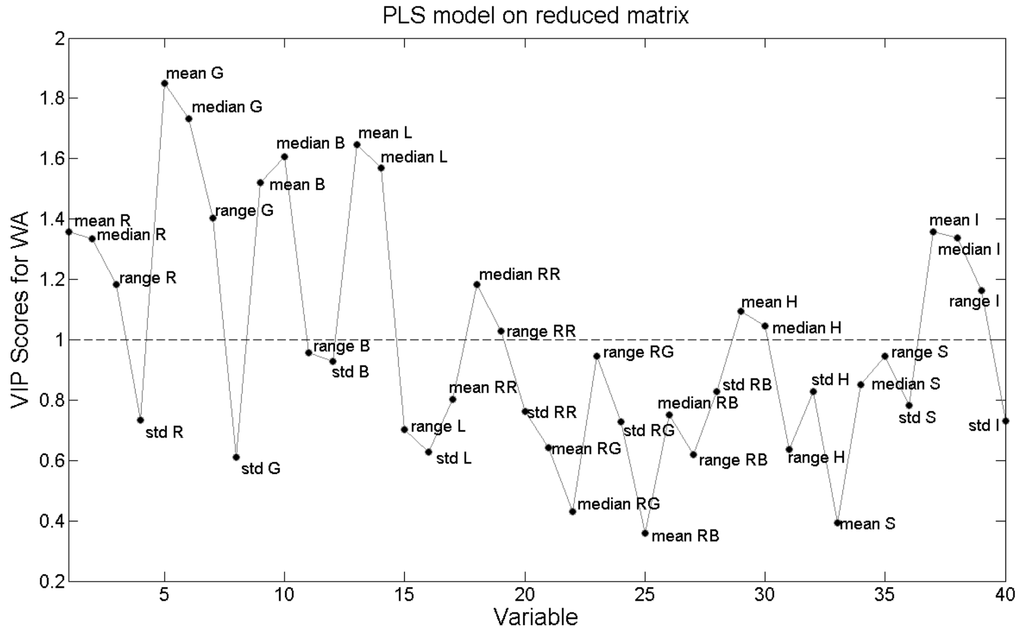

Figure 1

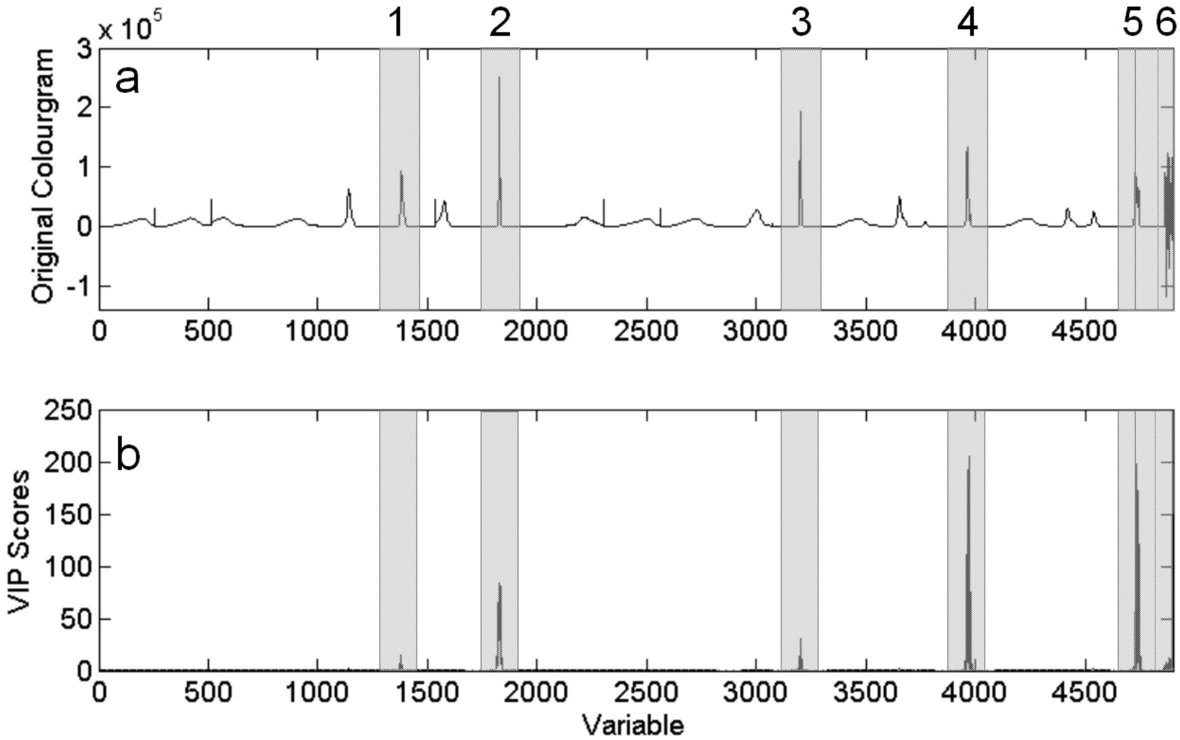
1
2
3
4
5
6
a
b
x 10^5
Original Colourgram
VIP Scores
Variable
0
500
1000
1500
2000
2500
3000
3500
4000
4500
3
2
1
0
-1
250
200
150
100
50

Figure 2

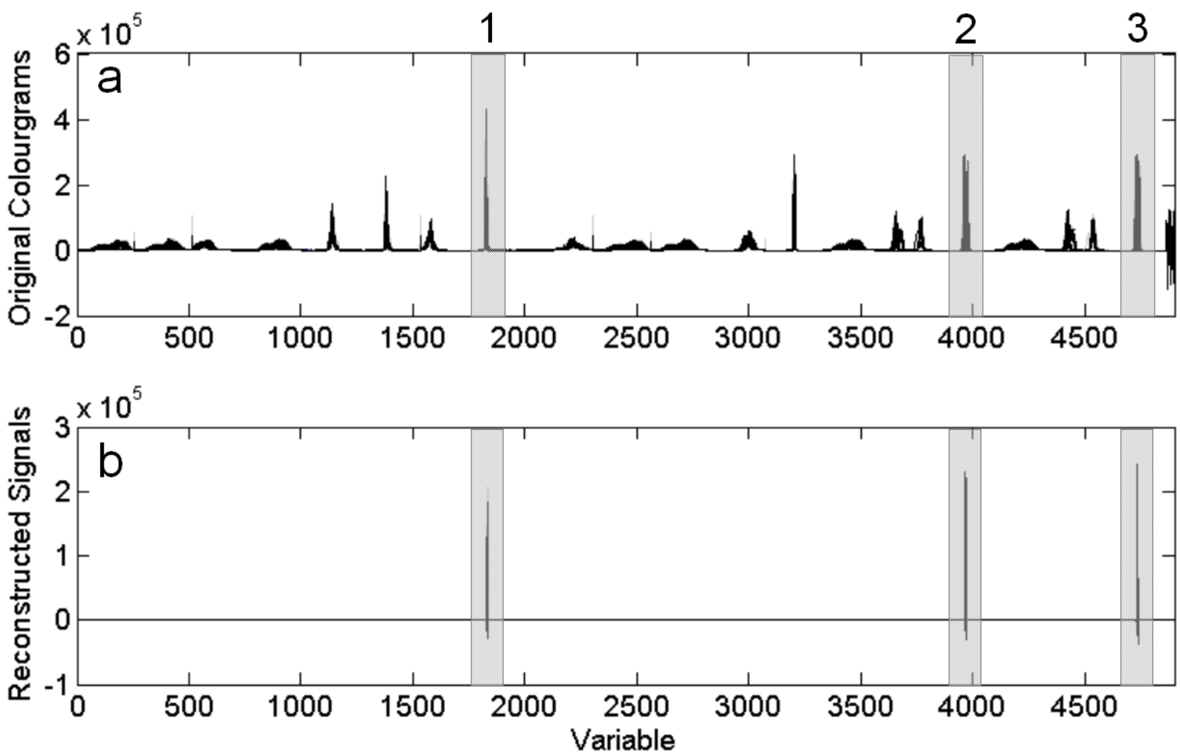
1
2
3
x 10^5
a
Original Colourgrams
0
500
1000
1500
2000
2500
3000
3500
4000
4500
6
4
2
0
-2
x 10^5
b
Reconstructed Signals
3
2
1
0
-1
Variable

Figure 3

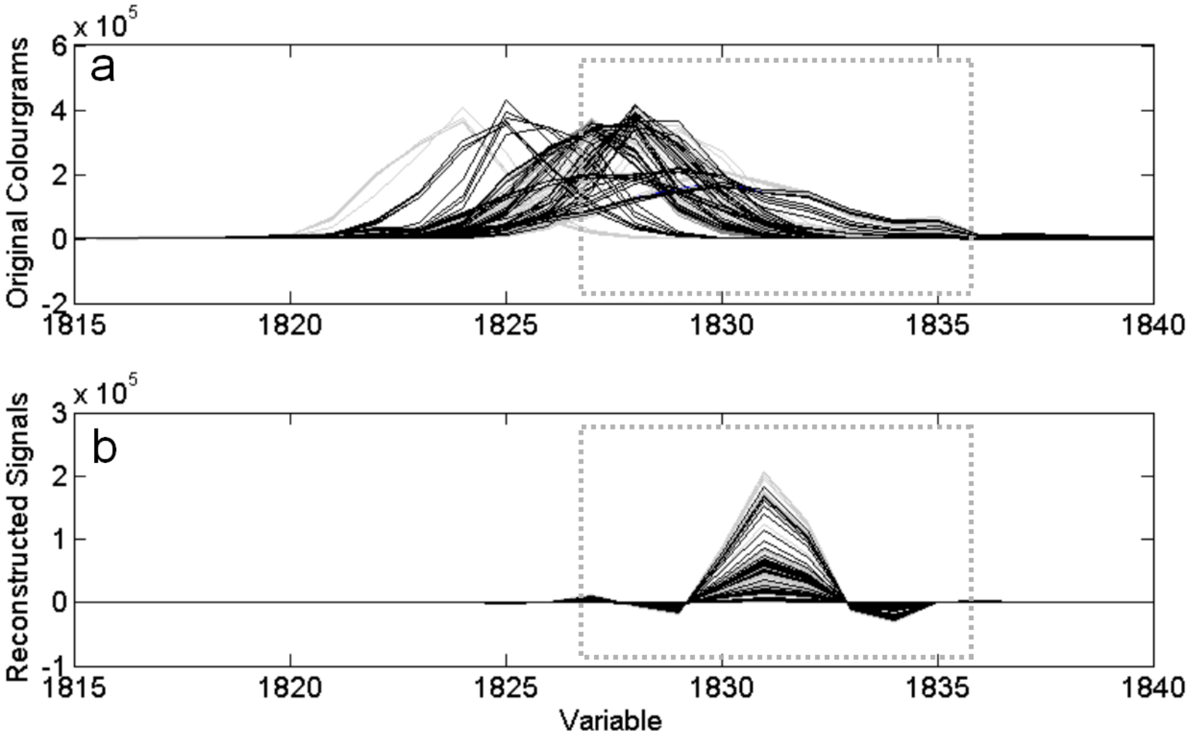
x 10^5
a
Original Colourgrams
6
4
2
0
-2
1815
1820
1825
1830
1835
1840
x 10^5
b
Reconstructed Signals
3
2
1
0
-1
1815
1820
1825
1830
1835
1840
Variable

Figure 4

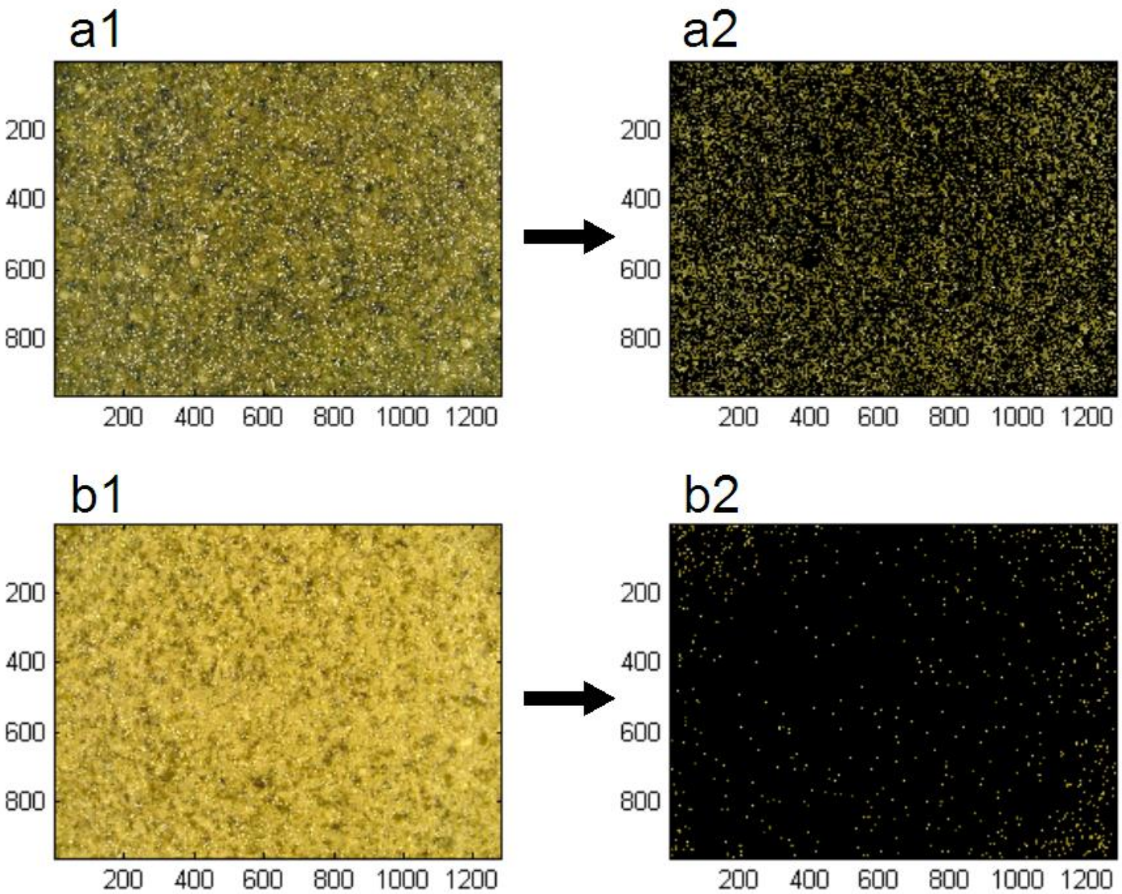

Figure 5