ELSEVIER

Chemometrics and Intelligent Laboratory Systems 62 (2002) 103–113

Chemometrics and intelligent laboratory systems

www.elsevier.com/locate/chemometrics

Simultaneous decomposition of multivariate images using three-way data analysis Application to the comparison of cereal grains by confocal laser scanning microscopy

Philippe Courcoux [a,*], Marie-Françoise Devaux [b], Brigitte Bouchet [b]

[a] *ENITIAA/INRA, Unité de Sensométrie et Chimiométrie, Domaine de la Géraudière BP 82225 44322 Nantes Cedex 03, France*
[b] *INRA URPOI, Rue de la Géraudière BP 71627 44316 Nantes Cedex 03, France*

Received 27 July 2001; accepted 30 January 2002

Abstract

We present a three-way data analysis method adapted to compare multispectral images such as those acquired by fluorescence microscopy. Spectral images are multivariate and can be considered as sets of pixels for which different spectral intensities have been measured. The algorithm proposed is based on the simultaneous decomposition of the covariance matrices calculated, in the present work, for fluorescence variables. This method allows the extraction of components common to each image and reveals their specificities by means of specific weights. The technique is illustrated by the analysis of microscopic image sequences acquired for five varieties of wheat and barley grains in 19 fluorescence conditions. The method made it possible to compare the fluorescence behaviors observed in the images and the strong similarities of the external tissues for the five cereals were highlighted.

Keywords: Multivariate image analysis; Confocal laser microscopy; Principal component analysis; Common components

1. Introduction

The development of computerised image-processing techniques makes it possible to describe objects using multivariate images. In the microscopy domain, acquiring multivariate images may be a means of identifying the different constituents or phases in complex samples. Such an approach can be developed in the domain of fluorescence microscopy. Fluorescence is a spectral phenomenon that occurs when light is absorbed by a substance and is re-emitted in the form of longer wavelength radiation. The fluorescence conditions are thus defined by a pair of wavelengths: the excitation (λ_{exc}) and emission wavelengths (λ_{em}). Fluorescence microscopy makes use of both naturally occurring fluorescent samples and specific fluorochrome dyes to visualise the components of interest. In multicomponent samples, each individual component is likely to exhibit a unique fluorescence behavior, though in a single image several components may fluoresce at the same time. The analysis of a sequence of spectral images, acquired by varying the fluores-

* Corresponding author. Tel.: +33-251-78-5436; fax: +33-251-78-5438.
E-mail address: courcoux@enitiaa-nantes.fr (P. Courcoux).

0169-7439/02/$ - see front matter
PII: S0169-7439(02)00013-8

cence spectral conditions λ_{exc} and λ_{em}, makes it possible to identify and segment the different constituents of a sample. This approach has been proved capable of labelling the biochemical constituents in cereal grain cell walls in which autofluorescent compounds such as lignin and phenolic acids are found [1,2].

Spectral images are basically multivariate and can be regarded as a three-way data structure. They can be interpreted both as a stack of images or as one image of pixel vectors. Considering pixels as vectors of spectral data independently of the spatial relationships occurring in the image scene is equivalent to unfolding each image into a single vector, and therefore the three-way multivariate image into a usual two-way data matrix [3]. Multivariate data analysis techniques can be applied to the resulting data matrix for the purposes of exploratory analysis, pattern recognition or segmentation. This image analysis approach is called multivariate image analysis (MIA), the principles of which are described in Geladi and Grahn [4]. In this context, principal component analysis (PCA) of three-way images may be considered as the basic technique [5]. Principal component regression (PCR) or partial least squares (PLS) regression can also be applied in order to relate features extracted from the stack of images to external information [6,7]. Classification and discrimination techniques have also been proposed to identify the homogeneous region in complex samples [8–10].

For the time being, most methods have been designed to analyse individual multivariate images and few papers relate the simultaneous analysis of several three-way images [11]. Four-way data structures are obtained when the time evolution of an object is observed or when different scenes have to be compared from multivariate images. In the present work, the objective was to compare the tissues observed in different cereal grains on the basis of their fluorescence properties as measured in microscopic images. The purpose was essentially descriptive and data treatments were sought to reveal the tissues common to different grains as well as to exhibit tissues specific to one or several samples. Common and specific fluorescent pixel behaviors, therefore, had to be noted from different sets of three-way images. Each image could be separately analysed by means of PCA, the components extracted being in some way related. This would turn the problem into the description of the relationships between the feature vectors in order to show the latent variables common to the different scenes analysed. Another approach would consist of concatenating all the images to be compared and applying a global PCA on the resulting concatenated array. This approach would have the advantage of directly exhibiting common fluorescent features but would not take into account the structure inside each scene.

Three-mode principal component analysis refers to a family of multivariate statistical techniques allowing the simultaneous analysis of several two-way data matrices [12]. In the context of analysing several multivariate images, the three modes refer to the pixels, the image being unfolded, the spectral conditions and the objects to be compared. Tucker, who introduced a procedure for analysing directly a three-way matrix, has given a general approach. Parameters are estimated for the three modes by means of alternating least squares. PARAFAC is a commonly used algorithm for fitting a Tucker model that assumes all the modes are common to all objects which is not generally the case with images. Pixels are not homologous when scenes have to be compared and the image size may vary from one scene to another. Alternatively, the parameters for the three modes can be extracted by simultaneously decomposing the cross-product matrices of each multivariate image. In the present work, we propose an algorithm that performs the decomposition into common components and weights specific to each scene.

2. Methods

2.1. Multi-way image analysis

Principal component analysis of a multidimensional image is described in order to introduce the principle and notations used in multivariate image analysis. Principal component analysis of a three-way ($I \times J \times M$) array of M ($I \times J$) images of the same scene leads to its decomposition into score matrices $\mathbf{T}_a$ and loading vectors $\mathbf{p}_a$. Using three-way matrix notations, one can express the multi-way image analysis of $\underline{\mathbf{X}}$ as

$$\underline{\mathbf{X}} = \sum_{a=1}^{A} \mathbf{T}_a \otimes \mathbf{p}_a + \underline{\mathbf{E}}$$

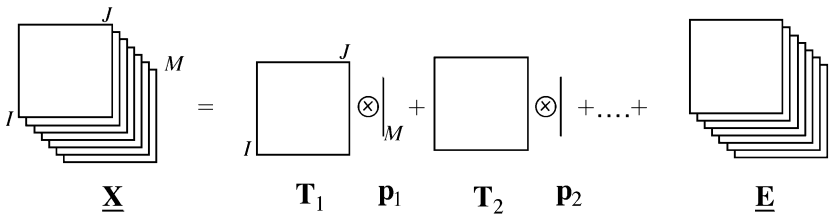


Fig. 1. Principal component analysis of a three-way table of M images of size $I \times J$. The decomposition results in score images $\mathbf{T}_a$ and loading vectors $\mathbf{p}_a$.

where $\underline{\mathbf{E}}$ is the residual of a rank-truncated PCA of $\underline{\mathbf{X}}$ and $\otimes$ denotes the Kronecker product.

This decomposition is illustrated in Fig. 1. The score matrices $\mathbf{T}_a$ have dimensions $I \times J$ and can be interpreted as images of the principal components and are ordered with a decreasing amount of inertia explained. The loadings $\mathbf{p}_a$ are vectors of size $M \times 1$ and express the contributions of the score images to the raw images.

Such an analysis reveals latent structures underlying the different images, the residual array $\underline{\mathbf{E}}$ being analyzed as the noise part of the decomposition.

This method is equivalent to the principal component analysis (PCA) of the unfolded matrix $\mathbf{X}$ derived from $\underline{\mathbf{X}}$, each row of the two-dimensional array $\mathbf{X}$ corresponding to one of the $(I \times J)$ pixels of the images while the columns refer to the M original variables (Fig. 2). The score vectors $\mathbf{t}_a$ issued from a PCA of $\mathbf{X}$ correspond to the unfolding of the score matrices $\boldsymbol{T}_a$.

$$\underline{\mathbf{X}} \overset{\text{unfolding}}{\longrightarrow} \mathbf{X} = \sum_{a=1}^{A} \mathbf{t}_a \mathbf{p}_a' + \mathbf{E}$$

Though the score vectors $\mathbf{t}_a$ can be reorganised as score images, it must be noticed that the analysis does not take into account the spatial position of the pixels.

The score vectors $\mathbf{t}_a$ are constrained to be orthogonal while the loading vectors $\mathbf{p}_a$ are orthonormal.

In practice, the dimensions of the matrix $\mathbf{X}$ make the components computation very time consuming when using the classical singular value decomposition (SVD) of $\mathbf{X}$. Alternatively, the loading vectors $\boldsymbol{p}_a$ can

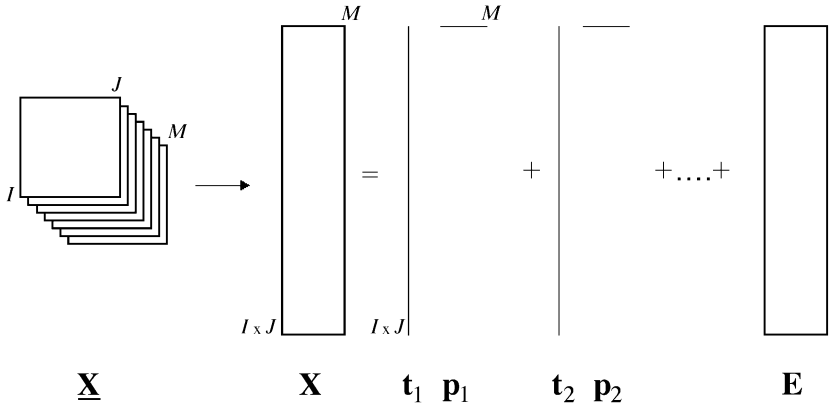


Fig. 2. Multivariate image analysis; the unfolded matrix $\mathbf{X}$ is decomposed using a principal component analysis into score images $\boldsymbol{t}_a$ and loading vectors $\mathbf{p}_a$.

be derived from the SVD of the cross-product matrix $\mathbf{X}'\mathbf{X}$:

$$\mathbf{V} = \mathbf{X}'\mathbf{X} = \mathbf{P}\Lambda\mathbf{P}'$$

where Λ is the diagonal matrix of eigenvalues. The score matrices are computed as $\mathbf{T} = \mathbf{X} \times \mathbf{P}$.

2.2. Analysis of several multivariate images

As a generalisation of the latter procedure, the analysis of several multivariate images can be developed as the simultaneous decomposition of several cross-product matrices. The objective is to reveal common latent components in the collection of multivariate images observed for different objects or scenes. In the present paper, the biochemical constituents common to several grain sections as well as specific ones are searched for according to their fluorescent behaviors.

The method is based on the assumption that common loading vectors exist that can possibly be weighted differently for each multivariate image. If we denote $\mathbf{X}^{(k)}$ the unfolded matrix derived from the k^{st} multivariate image, the model is

$$\mathbf{V}^{(k)} = \mathbf{X}^{(k)\prime}\mathbf{X}^{(k)} = \mathbf{P}\Lambda^{(k)}\mathbf{P}' = \sum_{a=1}^{A} \lambda_a^{(k)} \mathbf{p}_a \mathbf{p}_a'$$

where $\mathbf{P}'\mathbf{P} = \boldsymbol{I}_A$ and $\boldsymbol{\Lambda}^{(k)}$ is a diagonal matrix.

The vectors of $\mathbf{P}$ are the loadings common to the $\boldsymbol{M}$ multivariate arrays $\boldsymbol{X}^{(k)}$. The diagonal of the matrix $\mathbf{\Lambda}^{(k)}$ contains the specific weights $\lambda_a^{(k)}$ of the latent dimension a for the k^{st} stack of images. The greater a salience $\lambda_a^{(k)}$ is, the more the component is present in the multivariate image.

The parameters of this model can be estimated using a method described by Qannari et al. [13] and applied in the context of the sensory analysis. The algorithm allows the estimation of the common components and specific weights described in the INDSCAL model [14] decomposing simultaneously several inner-products matrices $\mathbf{X}^{(k)}\mathbf{X}^{(k)\prime}$. The same strategy can be applied to the cross-products $\mathbf{X}^{(k)\prime}\mathbf{X}^{(k)}$ to focus on the common and specific parts of a set of covariance matrices.

Table 1
Fluorescence conditions

Image number	Excitation wavelength (nm)	Emission wavelengths (nm)
1	633 (red laser)	>665
2	543 (green laser)	>570
3	543	575–640
4	543	>665
5	488 (blue laser)	>515
6	488	510–525
7	488	515–565
8	488	>570
9	488	575–640
10	488	RG 665
11	364 (UV laser)	>397
12	364	400–435
13	364	450–490
14	364	>515
15	364	510–525
16	364	515–565
17	364	>570
18	364	575–640
19	364	RG 665

In this case, the common loading vectors and the specific weights are defined in order to minimise the following loss function

$$L = \sum_{k=1}^{K} \left\| \mathbf{X}^{(k)\prime}\mathbf{X}^{(k)} - \sum_{a=1}^{A} \lambda_a^{(k)} \mathbf{p}_a \mathbf{p}_a' \right\|^2 \quad (1)$$

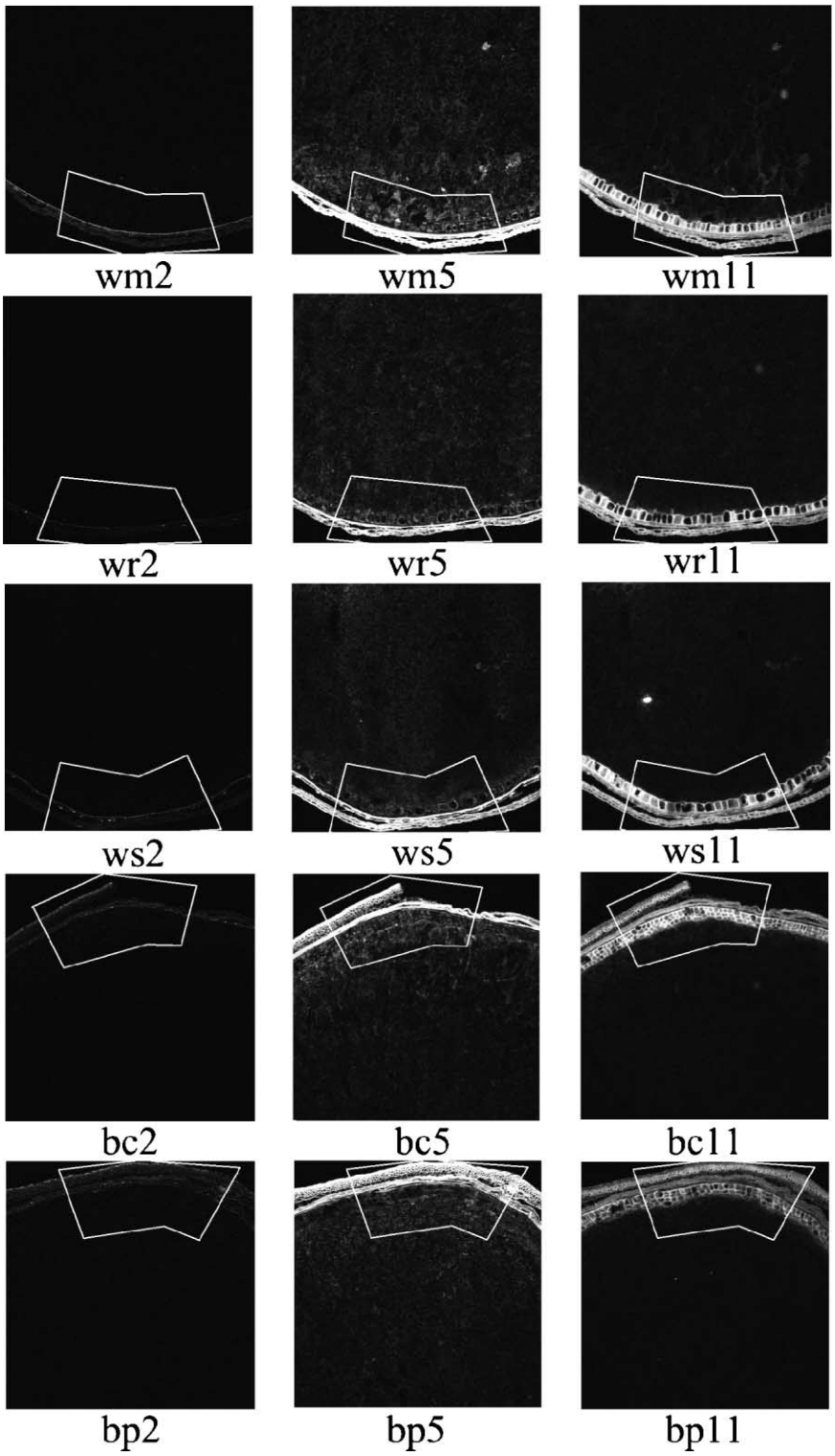
The algorithm solves this minimisation problem iteratively; the first, second, ..., sets of parameters corresponding to the different latent variables being estimated step by step. For a given dimension a, the common loading vector $\mathbf{p}_a$ and the specific weights $\lambda_a^{(k)}$ are computed using an alternate least squares technique.

At the first step, the loss function to be minimised is

$$L_1 = \sum_{k=1}^{K} \left\| \mathbf{X}^{(k)\prime}\mathbf{X}^{(k)} - \lambda_1^{(k)} \mathbf{p}_1 \mathbf{p}_1' \right\|^2$$

Assuming we have provisional estimates of the weights $\lambda_1^{(k)}$, the first common loading vector $\mathbf{p}_1$ is

Fig. 3. Examples of microscopic images for the five cereal grains: wm: wheat Market; wr: wheat Recital; ws: wheat Soissons; bc: barley Clarine; bp: barley Prisma. First column: green laser (image 2), second column: blue laser (image 5), third column: UV laser (image 11).


wm2
wm5
wm11
wr2
wr5
wr11
ws2
ws5
ws11
bc2
bc5
bc11
bp2
bp5
bp11

estimated by computing the first eigenvector of the weighted sum of the individual covariance matrices

$$\sum_{k=1}^{K} \lambda_1^{(k)} \mathbf{X}^{(k)\prime} \mathbf{X}^{(k)} \tag{2}$$

The weights minimising the loss function L_1 are then set by

$$\lambda_1^{(k)} = \mathbf{p}_1' \mathbf{X}^{(k)\prime} \mathbf{X}^{(k)} \mathbf{p}_1 \tag{3}$$

These estimations of $\mathbf{p}_1$ and $\lambda_1^{(k)}$ are repeated iteratively until no substantial change is observed in the loss function.

When convergence is reached for the first dimension, the procedure is repeated on the residual matrices $\tilde{\mathbf{X}}_2^{(k)} = \mathbf{X}^{(k)}(\mathbf{I} - \mathbf{p}_1 \mathbf{p}_1')$ to compute the second set of parameters.

The matrix of common loadings is constructed by merging the different vectors $\mathbf{p}_a$ in a matrix $\mathbf{P}$ and the computation of the score images is achieved by using the formula

$$\mathbf{T}^{(k)} = \mathbf{X}^{(k)} \mathbf{P} \tag{4}$$

The columns of $\mathbf{T}^{(k)}$ are the score images according to the common dimensions and can be visualized after reorganization. It is worth noting that the score images are not orthogonal as it is not possible to impose the simultaneous constraints of orthogonality on loadings and components when using three-way analysis.

An obvious advantage of the algorithm is that configurations are nested, i.e. the solution with A dimensions contains the solution with $A-1$ dimensions. The decreasing of the loss function can help the user to choose the number of dimensions to incorporate in his model.

This model can be compared to the common principal components model [15] which is a generalisation of principal components analysis to several groups of individuals. The main assumption is that the principal component transformations are identical in all the groups while the variances of these components may be specific to each group. The model can be interpreted as a simultaneous spectral decomposition of several covariance matrices. Flury and Gautshi [16] proposed an algorithm for the maximum likelihood estimation of the common principal components and the specific weights which is a generalisation of the Jacobi method for computing characteristic vectors and root of a symmetric matrix. The main difference with our approach is that the dimension of the solution has to be a priori chosen; therefore estimations with different dimensions are not nested and a test of the dimensionality of the solution has to be performed.

Another formulation of this model is the parallel factor analysis (PARAFAC). The algorithms are also based on alternate least squares but assume that the different arrays have the same individuals; in the image analysis context, as the different images do not necessarily have the same length, these techniques cannot be applied without modification.

3. Experimental

3.1. Samples

Cereal grains from three wheat varieties (Recital, Soissons and Market) and two barley varieties (Clarine and Prisma) were studied. For each sample, one grain was cut after hydration for 4 h on wet paper in a petri box. Transverse sections of 10 μm thick were cut in the middle of the grains.

3.2. Multivariate images

Images of the natural fluorescence occurring in the samples were acquired using a confocal laser scanning microscope Zeiss LSM410. The microscope is equipped with four lasers providing four excitation wavelengths: 364 (ultraviolet light), 488 (blue light), 543 (green light) and 633 nm (red light). The light emitted by fluorescence was collected through nine filters being either long pass (i.e. collecting all the light emitted over a certain wavelength) or band pass (i.e. designed for collecting the light emitted between two given wavelengths). As fluorescence only occurs at wavelengths higher than the excitation wavelength, the microscope can collect sequences of 19 fluorescence images. The precise description of the fluorescence images is given in Table 1. For example, image 1 was obtained by lighting samples with excitation wavelength 633 nm and recording the light emitted over 665 nm, using the long pass filter “>655”. Image

Table 2
Loss function and percentage of variance explained by the first dimensions

Dimension	Loss function	Percentage of variance explained
1	0.32	93.60
2	0.108	4.23
3	0.101	0.15
4	0.101	0.01

6 was obtained by lighting samples with excitation wavelength 488 nm and recording the light emitted between the two stopping wavelengths of the band pass filter: 510 and 525 nm. The multivariate images were acquired with the $\times 10$ magnification objective allowing the visualisation of about one sixth of the grain. Each individual image was digitised as 512×512 pixel tables for which the fluorescence intensities were coded with 256 grey level values from 0 (black) to 255 (white). Under these conditions the images were 1270×1270 μm large, one pixel corresponding to 2.48×2.48 μm. In order to compare the fluorescence intensities of the five multivariate images, the contrast and brightness of the images were fixed at the same level for each excitation laser and for the five grains studied. The dimensions of the resulting data set were 5 grains $\times$ 512 lines $\times$ 512 columns $\times$ 19 fluorescence conditions.

3.3. Selection of regions of interest

In order to reduce the amount of data for calculating the model of decomposition of the four-way data table, regions of interest were selected in the images. The regions had been chosen by the operator in order to select pixels from each layers of the grain, including a part of the starchy albumen, which was not

Table 3
Weights for the five tables on the first three dimensions

Grain	Dim 1	Dim 2	Dim 3
Wm	0.968	0.188	0.049
Wr	0.972	0.202	0.033
Ws	0.958	0.244	0.032
Bc	0.967	0.248	0.037
Bp	0.971	0.119	0.038

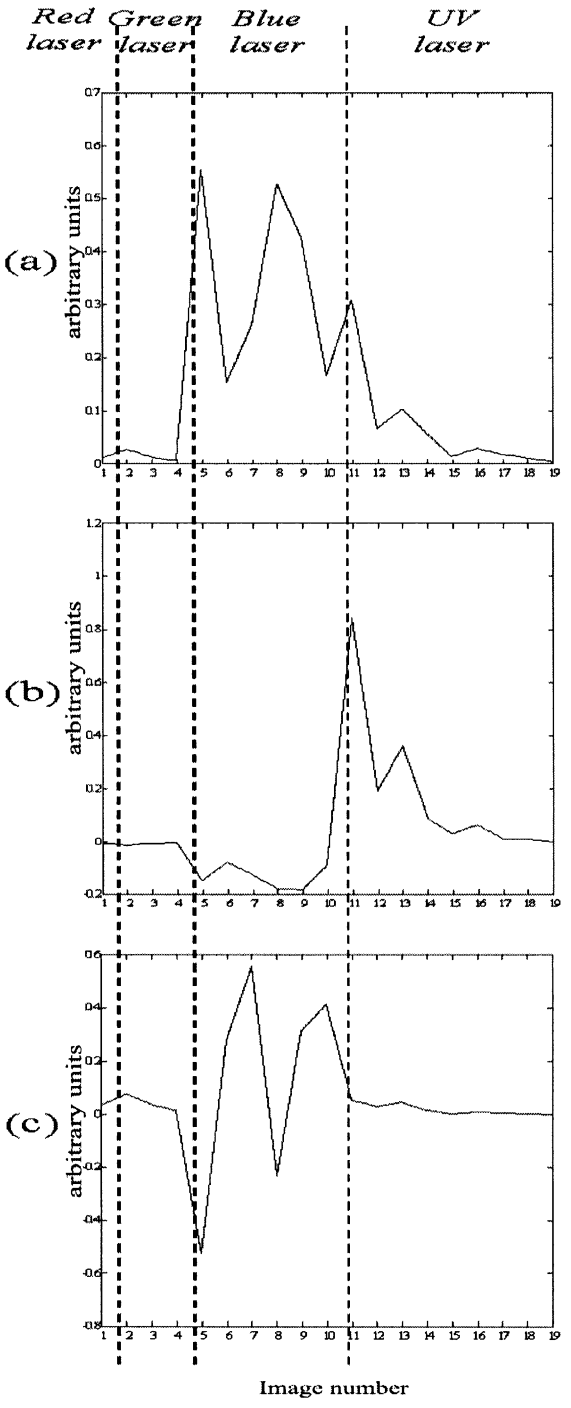


Fig. 4. Loadings of the 19 images on the first three dimensions (a, b and c, respectively).

fluorescent, and a part of the slides around the section (Fig. 3). This resulted in slightly different sizes of data table for the five grains varying from 34615 to 43185 pixels.

3.4. Application of common components and specific weights analysis

Each individual data table $\mathbf{X}^{(k)}$ was standardized prior to the analysis so that the norms of covariances matrices $\|\mathbf{X}^{(k)'}\mathbf{X}^{(k)}\|^2$ were equal to 1. This pretreatment was performed to take into account the different numbers of pixels and the uncontrolled fluorescence intensity variations that can be encountered during image acquisition.

The model of decomposition was assessed from the data collected for the regions of interest from Eqs. (2) and (3). In order to calculate the score images for the complete scene rather than for the region of interest, Eq. (4) was applied to all the pixels of the images.

3.5. Results and discussion

Fig. 3 shows the different tissues that could be observed from their natural fluorescence. Using the green laser, only single layers were fluorescent in wheat market, barleys Clarine and Prisma, which can be identified as waxy layers for barley and as the seminal tegument for wheat. After excitation using the blue laser, all the external tissues were highlighted corresponding to pericarps for wheat and to husks (the most external layer) and pericarps for barley. The waxy layers were also fluorescent under these conditions. The biochemical compounds responsible for this fluorescence are still to be found [17]. After illumination with the UV laser, an additional tissue corresponding to the aleurone layer became fluorescent. This tissue consists of one layer of large cells for wheat and three layers of small cells for barley. Ferulic acid is known to be responsible for this fluorescence. When using the red laser, almost no fluorescence could be seen (data not shown). The images in Fig. 3 reveal that several tissues can fluoresce in a single image, though their fluorescence behavior may be specific. It was expected that appropriate multivariate image data treatments would make it possible to identify individually the different fluorescence behaviors observed among the images and therefore provide a tool to highlight the similarities and specificities of each cereal studied. The technique of common components and specific weights analysis was applied for this purpose.

Table 2 reports the evolution of the loss function for the first four dimensions. This loss function quantifies the residual variance after each estimation step, cumulated for all the image data tables. The total initial variance, measured by the sum of the norms of $\mathbf{X}^{(k)'}\mathbf{X}^{(k)}$, was equal to 5. The decreasing of the loss function expresses the part of this quantity which is explained by each dimension. In this case, the first three dimensions explain 97.98% of the total variance. These first three dimensions were considered sufficient to reveal the structure of the original images.

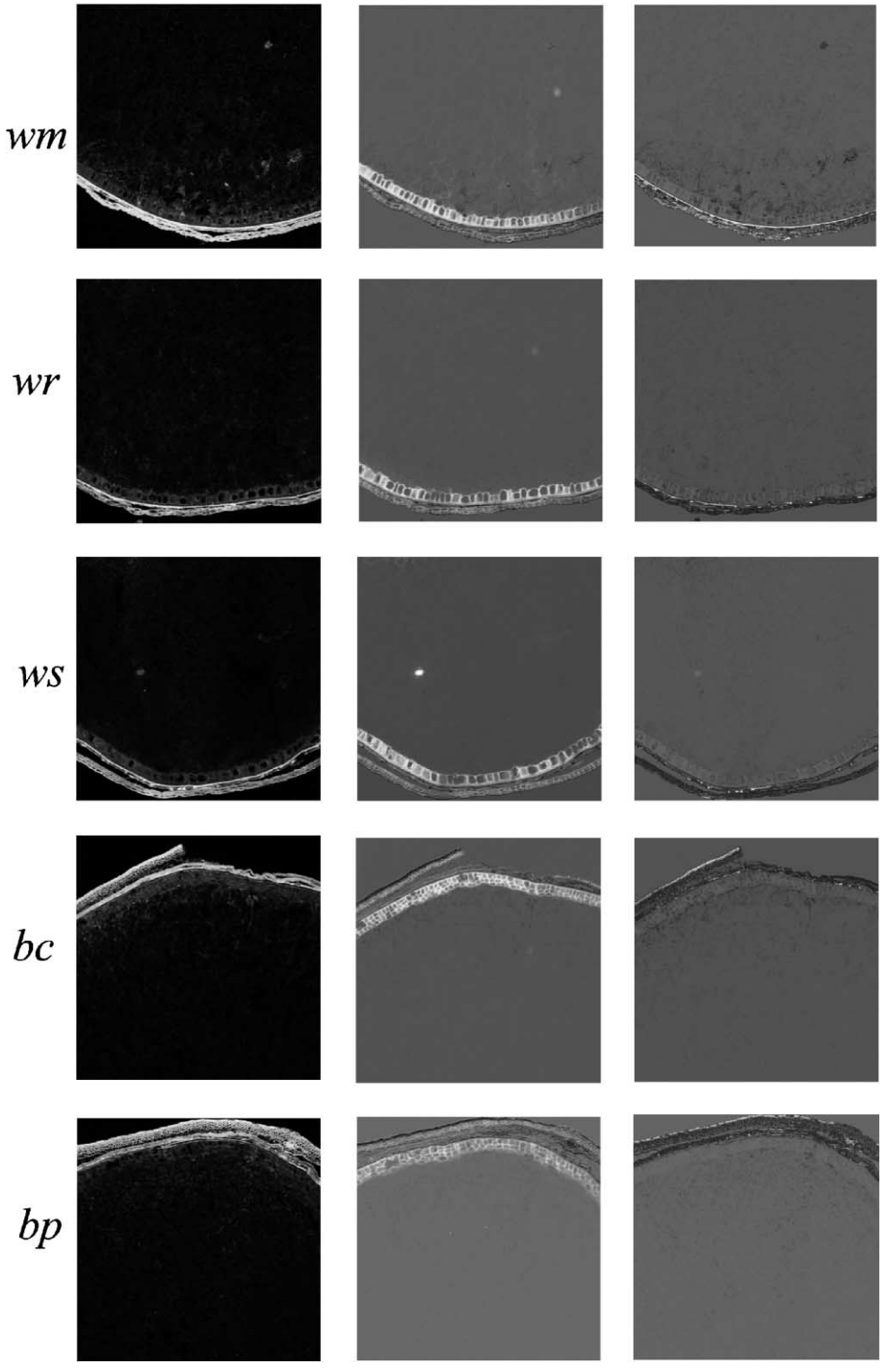
The weights of the five types of grains for the first three dimensions are presented in Table 3. They exhibit the relative abundance of the common structures in the original images.

The weight $\lambda_a^{(k)}$ corresponds to the variance of the image score $\mathbf{t}_a^{(k)}$:

$$\mathbf{t}_a^{(k)'}\mathbf{t}_a^{(k)} = \mathbf{p}_a'\mathbf{X}^{(k)'}\mathbf{X}^{(k)}\mathbf{p}_a = \lambda_a^{(k)}$$

They can be used as a diagnostic tool to compare samples or to detect an abnormal scene. Comparing different weights on a dimension gives a look on similarities between samples. These similarities between objects will depend on the presence of pixels having the same behavior of fluorescence and their relative amounts in the different images. For this application, the definition of the regions of interest (their positions, sizes and shapes) will modify these relative proportions and consequently the values of the weights. For the selected regions of interest, the weights (Table 3) are very homogeneous for the five cereals, showing very similar fluorescence behaviors.

Fig. 5. Score images on the first three dimensions. The three columns refer to the dimensions 1, 2 and 3. wm: wheat Market; wr: wheat Recital; ws: wheat Soissons; bc: barley Clarine; bp: barley Prisma.


wm
wr
ws
bc
bp

There is also a close link between the loss function and the weights: at each stage, the decreasing of the loss function is equal to the sum of the squared weights.

$$L = \sum_{k=1}^{K} \| \mathbf{X}^{(k)\prime}\mathbf{X}^{(k)} - \sum_{a=1}^{A} \lambda_a^{(k)} \mathbf{p}_a \mathbf{p}_a' \|^2$$

$$= \sum_{k=1}^{K} \| \mathbf{X}^{(k)\prime}\mathbf{X}^{(k)} \|^2 - \sum_{k=1}^{K} \sum_{a=1}^{A} \left(\lambda_a^{(k)}\right)^2$$

As each dimension a describes a part of the initial inertia, the importance of this dimension may be assessed by the percentage of inertia explained by the component a using the weights $\lambda_a^{(k)}$

$$I_a = \frac{\sum_{k=1}^{K} \left(\lambda_a^{(k)}\right)^2}{\sum_{k=1}^{K} \| \mathbf{X}^{(k)\prime}\mathbf{X}^{(k)} \|^2}$$

Fig. 4 shows the loading vectors computed for the first three dimensions. The components of loading vector $\mathbf{p}_a$ are the coefficients of the linear combinations used to compute the score images for dimension a. They exhibit the common features of the raw data tables, the way in which this part is common being given by the $\lambda_a^{(k)}$; in our case, the loadings express the common fluorescence behaviors.

On the first dimension, the highest coefficients were observed for images acquired after excitation with the blue laser, and to a lesser extent with the UV laser, corresponding to images with high levels of fluorescence. The first dimension could therefore be interpreted as indicating the raw images for which some fluorescence was observed. The second loading vector revealed a contrast between the fluorescence behaviors after excitation with the blue and UV lasers. The third loading showed variations observed after excitation with the blue laser alone between the long pass filters >515 and >570 nm and the other filters.

The model of decomposition of the multivariate images assessed from the region of interest was applied to all the pixels of the images. The score images were reconstituted by folding the score vectors and the first three dimensions are shown in Fig. 5. In these images, the grey level intensities were arbitrary and set independently for all the images. As a consequence, they must be interpreted as revealing the tissues contrasted by the method, i.e. black and white corresponding respectively to the negative and positive coefficients of the loading vectors.

The first score images were very similar to the images obtained after excitation with the blue laser, the aleurone layer being slightly visible in grey. As indicated by the first loading vector, the areas for which fluorescence occurred were highlighted in the score images. Thus, the pericarp, husks and waxy layers were overall the most fluorescent tissues considering all the conditions used in this experiment. The second score images mainly exhibited the aleurone layers in white for the five cereals as expected from the UV images and the second loading vector. In addition, the husks were found to have a similar fluorescent behavior to the aleurone layer. The pericarp was slightly white for the three wheat samples and almost grey in the barley images. The external waxy layers were found contrasted in black corresponding to their intense fluorescence observed after excitation with the blue laser. The fluorescence in the UV range is known to be attributed to ferulic acids [18]. Regarding the second score images, ferulic acid could be found in the aleurone layer, the husk and the wheat pericarp. The waxy external layers were highlighted in white for the two barley samples in the third score images together with the seminal tegument of wheat Market, in contrast with the pericarps and husks in black. This contrast was obtained by considering the fluorescence using the blue laser light. The score images showed strong similarities in the tissue of the five grains studied. Three main fluorescence behaviors can be pointed out found in (i) the aleurone layer and the husks, (ii) the pericarps and the husks, (iii) the waxy layers and seminal tegument. As already mentioned, the fluorescence of the aleurone layers is caused by ferulic acid. Fincher and Stone [19] have reported some studies indicating the presence of phenolic acids in barley husks that can explain the fluorescence found in husks. The third score images revealed the fluorescence common to pericarp and husk for barley and wheat. The differences observed between the wheat and barley pericarp in the second score images cannot be interpreted straightforwardly because only relative grey levels are represented on the images. The third score images also revealed the

fluorescence common to the external waxy layers and the seminal tegument, the trace of which is found in wheats Soissons and Recital. The study therefore shows that at least three fluorescent compounds are responsible for autofluorescence in wheat and barley, the score images showing their localisation in situ in the grains.

4. Conclusion

The simultaneous decomposition of cross-product matrices using common component and specific weight analysis made it possible to compare different stacks of fluorescence images obtained by varying the conditions of acquisition. The method was applied to the comparison of wheat and barley varieties and made it possible to identify the different fluorescence behaviors in the global set of images and to reveal the localisation of common fluorescence in the five cereals studied.

The procedure for extracting the common and specific parts of the images is based on an alternating least square algorithm that can be seen as a simple way to achieve simultaneous principal component analyses on different multivariate images. The model proposed allows the original data tables to be of different sizes, which is generally the case in image analysis. This aspect is an obvious advantage over the PARAFAC algorithm. In addition, the fact that solutions are nested leads to a relatively easy interpretation of the decomposition. The examination of the decreasing of the loss function provides the practitioner with a useful tool to discard the dimensions reflecting noise in a similar way than to the examination of the decreasing of eigenvalues in principal component analysis. As a consequence, the decomposition by common component and specific weights analysis also acts as a data reduction technique.

References

[1] P.M. Baldwin, D. Bertrand, B. Novales, B. Bouchet, G. Collobert, D.J. Gallant, Anal. Chem. 69 (1997) 64339–64349.
[2] B. Novales, D. Bertrand, M.F. Devaux, P. Robert, A. Sire, J. Food Sci. Agric. 71 (1996) 376–392.
[3] K. Esbensen, P. Geladi, Chemom. Intell. Lab. Syst. 7 (1989) 67–86.
[4] P. Geladi, H. Grahn, Multivariate Image Analysis, Wiley, New York, 1996.
[5] P. Geladi, H. Isaksson, L. Lindqvist, S. Wold, K. Esbensen, Chemom. Intell. Lab. Syst. 5 (1989) 209–220.
[6] K. Esbensen, P. Geladi, H.F. Grahn, Chemom. Intell. Lab. Syst. 14 (1992) 357–374.
[7] 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, Chemom. Intell. Lab. Syst. 35 (1996) 187–197.
[8] P. Geladi, J. Serts, F. Lindgren, Chemom. Intell. Lab. Syst. 24 (1994) 145–167.
[9] P. Robert, D. Bertrand, M.F. Devaux, A. Sire, Anal. Chem. 64 (1992) 664–667.
[10] D. Bertrand, B. Novales, Y. Chtioui, SPIE International Symposium, 1–6 November 1998, Boston, USA, 1998.
[11] N. Gouti, P. van Espen, M.H. Feinberg, Chemom. Intell. Lab. Syst. 47 (1999) 21–31.
[12] P.M. Kroonenberg, Three-mode Principal Component Analysis, DSWO Press, Leiden, The Netherlands, 1983.
[13] E.M. Qannari, I. Wakeling, P. Courcoux, H.J.H. MacFie, Food Qual. Preference 11 (2000) 151–154.
[14] J.D. Caroll, J.J. Chang, Psychometrika 40 (1970) 33–51.
[15] B. Flury, Common Principal Components and Related Multivariate Models, Wiley, New York, USA, 1988.
[16] B. Flury, W. Gautschi, SIAM J. Sci. Stat. Comput. 7 (1986) 169–184.
[17] S.J. Symons, J.E. Dexter, J. Cereal Sci. 23 (1996) 73–83.
[18] A. Saadi, I. Lempereur, S. Sharonov, J.C. Autran, M. Manfait, J. Cereal Sci. 28 (1998) 107–114.
[19] G.B. Fincher, B.A. Stone, Adv. Cereal Sci. Technol. VIII (1986) 207–293.