ELSEVIER

Chemometrics and Intelligent Laboratory Systems 42 (1998) 141–158

Chemometrics and intelligent laboratory systems

Multivariate feature extraction from textural images of bread

Knut Kvaal *, Jens Petter Wold, Ulf G. Indahl, Pernille Baardseth, Tormod Næs

MATFORSK, Osloveien 1, N-1430 Ås, Norway

Abstract

In order to compute the classical texture measures there is often a need to perform extensive calculations on the images and do a preprocessing in a specialised manner. Some of these texture measures are constructed to estimate specific information. Other texture measures seem to be more global in nature. The techniques presented in this paper define algorithms applied on the raw image without extensive preprocessing. We want to show that mathematical transformations of images on a vectorised form will easily enable the use of multivariate techniques and possibly model several features hidden in the images at the same time. In this paper we will compare five different methods of extracting features from textural images in food by multivariate modelling of the sensory porosity of wheat baguettes. The sample images are recorded from factorial designed baking experiments on wheat baguettes. The multivariate feature extraction methods to be treated are the angle measure technique (AMT), the singular value decomposition (SVD), the autocorrelation and autocovariance functions (ACF) and the so-called size and distance distribution (SDD) method. The methods will be tested on equal basis and the modelling of sensory porosity from extracted features is done using principal component regression (PCR) and partial least square regression (PLS). The difference between the behaviour of the methods will be discussed. The results show that all the methods are suited to extract sensory porosity but the AMT method prove to be the best in this case.

Keywords: Image analysis; Texture; Food; Cereals; Baguettes; Feature extraction; Multivariate analysis; PCR; PLS; ACOR; ACOV; AMT; SVD; SSD

1. Introduction

The traditional use of sensory analysis to measure food quality has shown to be very useful in many practical cases [1,2], but it is both time-consuming and expensive. This has motivated the development of alternative techniques for more rapid and efficient measurement of key characteristics of the products. Image analysis is one such interesting technique. Applications of image analysis in measurement of food quality have already shown to be a promising area of research [3–5], and are expected to be of significant practical importance in the coming years. The focus of the present paper will be on a specific area of image analysis—analysis of textures—and its relationship to sensory measurements.

Traditionally, analysis of textural images has been based on counting objects, measuring object areas and on combining such information in different ways [6,7]. These approaches are closely linked to a physical understanding of the texture measures and are often handled by univariate techniques. Recent research has, however, focused more strongly on ex-

* Corresponding author. Tel.: +47-64-97-01-00; fax: +47-64-97-03-33; e-mail: knut.kvaal@matforsk.no.

0169-7439/98/$19.00
PII: S0169-7439(98)00017-3

tracting features by analysing the pixel distributions directly without any special attention given to the objects in the image. Such approaches are typically multivariate in the sense that they extract a whole vector of features from each image [8–10]. Multivariate statistical analysis is therefore required in order to understand the relationship among the variables and to find relationships to external sensory variables.

In the present paper we will present and compare a number of alternative multivariate feature extraction methods. Most of the methods have been published and discussed elsewhere, but in some cases we make modifications to make them more suitable for the present application. The comparison of the methods will focus on their ability to predict externally measured (sensory) texture variables. The images used in the study are recorded from wheat bread baguettes produced from different flour varieties and by using different baking conditions. Although there is a manifold of articles in the literature on analysis of texture images, there seems to be very few comparisons of this kind. Our goal is also to show that images taken in robust environments when it comes to lighting conditions should be possible to model with little or no preprocessing.

The multivariate feature extraction methods to be treated below are: the angle measure technique (AMT), the singular value decomposition (SVD), the autocorrelation and autocovariance functions (ACF) and the so-called size and distance distribution (SDD) method. The SVD and AMT methods have already shown very promising results in the characterisation of sensory attributes [4,11]. The ACF method has been proposed elsewhere for texture analysis, but seems to have obtained little attention in the literature, especially so in the more applied area. The SDD method is a new and multivariate analogue to established methods for analysing object sizes and their distribution. Instead of looking at either the sizes or distances separately, the method brings both aspects into consideration simultaneously. Although not necessary for computing any of the above mentioned feature vectors, all the methods are most naturally introduced by assuming the images are isotropic and stationary [12]. To have an objective comparison between the different feature extraction methods we have decided not to optimise the single methods.

The relationship between the feature vectors obtained by the four different methods and the sensory attribute porosity will be obtained by the linear regression methods PCR and PLS. No attempts in the direction of nonlinear modelling will be tried. The PLS and PCR are data compression methods especially useful for collinear spectral data which will be a problem for all the feature extraction methods considered here.

2. Experimental

2.1. Design, materials and methods

2.1.1. Dough preparation and processing

Wheat baguettes were produced with a defined recipe described in Ref. [13]. An experimental design with four factors was used. The factors were flour quality (four levels), mixing time (two levels), proofing time (two levels) and baking process (two levels). The levels were combined in a full factorial experiment (32 experiments). A total of 10 baguettes were produced for each of the 32 combinations of factors. Each baguette was divided in two. Half portions from each of the 10 baguettes were sent to sensory analysis. Ten sensory assessors got one-half each. The average over assessors was used as the reference value. Since the present paper is a study of feature extraction methods, we decided to concentrate on only one of the sensory attributes, namely porosity. Three of the 10 baguettes were used for image analysis. This gave three independent images (three replicated measurements) for each of the 32 combinations of factors. In other words, 96 images were recorded. An overview of the experiment is given in Fig. 1.

2.1.2. Sensory reference values

Baguettes were analysed by conventional sensory profiling using 10 trained panellists. Members were selected and trained according to guidelines in ISO 6564:1985-E and ISO/DIS 8586-1:1989. Panellists developed a lexicon by describing differences between extreme samples and developed a consensus list of 13 sensory attributes for profiling. These were glossiness, brittles of the crust, porosity, surface of characteristic cuts, elasticity, intensive odour, fresh smell, intensive flavour, fresh flavour, salt flavour,

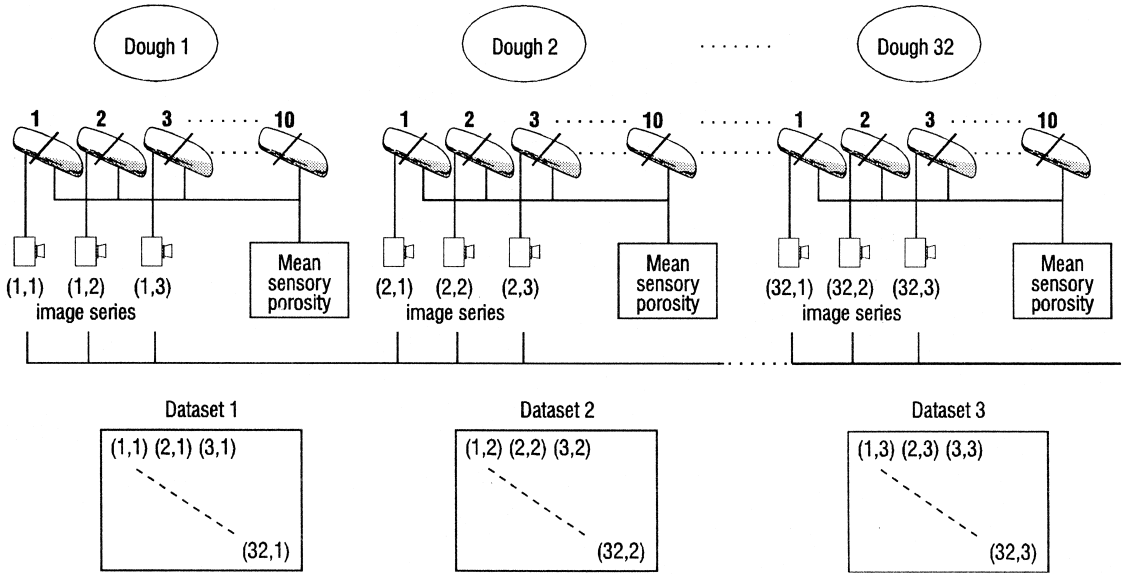

Fig. 1. Construction of image series used in calculations.

firmness, juiciness and crispness of the crust. Panellists were trained in use of definitions of sensory attributes and rating anchors by pretesting extreme samples. Each panellist was given one-half baguette on a white dish. A continuous nonstructured scale was used for evaluation. The left side of the scale corresponded to the lowest intensity (value 1.0) and the right side corresponded to the highest intensity (value 9.0). Evaluations were done in a laboratory equipped as described in ISO 8589:1988-E. Each panellist evaluated the samples at individual speed on a computerised system for direct recording of data (CSA Compusense, version 4.2, Canada). Samples were served in randomised order. No replicates were done. In this paper we use only the sensory porosity as reference value. Thus for each replicate block of images, there are the same sensory porosity values (the mean over assessors, Fig. 1).

2.2. Image measurements

The images were produced using a Canon EX2 video camera and recorded with the ScreenMachine II frame grabber. The signal from our camera was composed of separated luminance and chrominance signals (Y/C). Experimentation with the light conditions led us to use 45° illumination from both sides. The objects were illuminated using four tungsten lamps, two from each side.

The recorded images were recorded in true colour at 512×512 pixel resolution and converted to 256 level grey scale. The images of the baguettes were all centred using a thresholding technique along with calculations of the centre of gravity of the image, a central rectangular cut-out and a resizing to obtain a 256×256 greyscale image. The images were enhanced with a 3×3 unsharp mask. The unsharp masking is the general process of subtracting a blurred image from an original [14]. The illuminations of the samples introduced slight light gradients due to the 45° illuminations from both sides. Introduction of a light gradient is often needed to enhance the pore structure. With these controlled lighting conditions, however, the levels of the mean pixel values of the grey level images may vary slightly due to different colours of the flour being used and also the treatment by the baking process. The samples had in fact different lightness which led to small differences in mean pixel intensity. The images could have been normalised in some way but was not done. Typical representative images of different textures are shown in Fig. 2. The samples are selected to span the porosity space. Samples 5, 15, 26 and 12 ranges from low to high degree of porosity. The corresponding

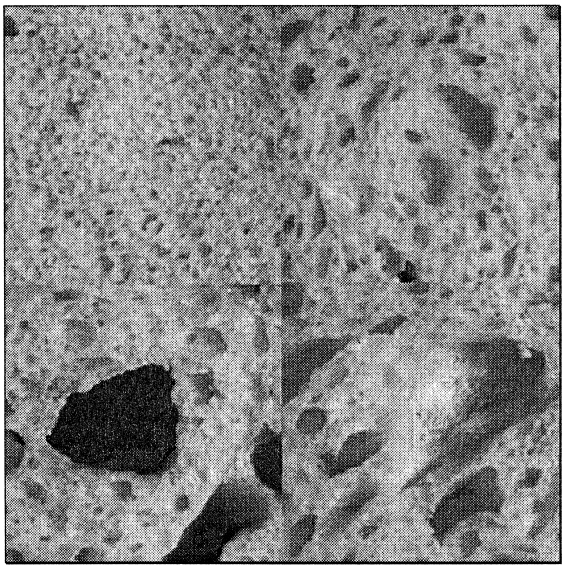

Fig. 2. Representation of four images with different porosity (samples 5, 15, 26 and 12 clockwise from upper left) recorded for one datablock.

images will be given special attention for the different feature extraction methods discussed later in this context.

2.3. Selection of calibration / validation sets—calculations

In this paper, no attempt was made at treating the 32 samples as representatives from a population of samples. Instead the 32 combinations of factors were themselves treated as the actual population of interest. The sensory reference value (average over assessors) was considered the true porosity value for each of the factor combinations. The present study is primarily a comparison of alternative feature extraction methods in their ability to provide valid information about this porosity. Therefore, the following cross-validation procedure was used for testing validity of the feature extraction methods. Each of the three independent image replicates were treated as three independent measurement series of the 'population' of interest. Each of the three independent series were treated as a cross-validation segment. In each cross-validation run (3 in total), 64 samples were used for calibration and 32 for testing. In this way, each replicate was treated in two calibration segments and one test segment. Note that when this strategy is used, no biological variation is part of the cross-validation, only variability among the replicated images. In other words, the variance in each RMSEP reported is due to measurement noise in $\boldsymbol{X}$ (the feature vector) only.

The calculations were carried out using Refs. [15,16] by Camo A/S and Watcom C++ ™.

3. Methods

3.1. Singular value decomposition (SVD)

For a technical description of the SVD algorithm we refer to Ref. [17]. In our approach the image is considered as a matrix of pixels ordered in rows and columns.

Consider the image $\mathbf{A}$ of size $(m \times n)$. The SVD theorem states that there exists unitary orthogonal matrices $\mathbf{U}$ and $\mathbf{V}$ of size $(m \times r)$ and $(n \times r)$, respectively, and a diagonal matrix $\mathbf{S}$ of size $(r \times r)$ (where r is the rank of $\mathbf{A}$) such that

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}'. \tag{1}$$

The matrix $\mathbf{S} = \{s_{ij}\}$ is considered to be a generalised spectrum of the image (Kvaal et al., 1996 [4]). The matrix $\mathbf{S}$ can be written as

$$S = \begin{pmatrix} s_{1,1} & 0 & 0 & 0 & 0 \\ 0 & .. & 0 & 0 & 0 \\ 0 & 0 & .. & 0 & 0 \\ 0 & 0 & 0 & .. & 0 \\ 0 & 0 & 0 & 0 & s_{r,r} \end{pmatrix}. \tag{2}$$

The singular values are sorted in descending order. By applying the SVD on the image $\mathbf{A}$ this is analogous to estimating the principal components of an arbitrary data-matrix, using the rows and columns of $\mathbf{A}$ as objects and variables respectively. The singular value (SVD) spectrum λ of an image is by definition the vector of diagonal elements from $\mathbf{S}$:

$$\boldsymbol{\lambda} = \mathrm{diag}(\mathbf{S}). \tag{3}$$

Here we will use a limited number of factors to model the actual sensory porosity, thus use a truncated version of the SV-spectrum:

$$\boldsymbol{\lambda}_p = \mathrm{diag}(\mathbf{S}_p)(\ p \le \min(m,n)) \tag{6}$$

The SV-spectra λ_{pi} $(1 \le i \le k)$ estimated using p factors from a set of k images can be arranged row-wise in the matrix $\mathbf{\Lambda}_p$ containing the transposed vectors λ'_{pi} as its rows:

$$\mathbf{\Lambda}_p = \left[\boldsymbol{\lambda}_{p1}, \boldsymbol{\lambda}_{p2}, \boldsymbol{\lambda}_{p3}, \ldots \boldsymbol{\lambda}_{pk} \right]'. \quad (7)$$

The singular values of images representing bread slices in this case have been reported to contain information of the image texture [4].

3.1.1. Implementation

The SVD was applied to each of the 96 images. Fig. 3 shows examples of different SV-spectra for the four selected images shown in Fig. 2. These are images selected to represent low, medium and high sensory porosity. Image 12 represents big extreme holes. This is shown in Fig. 3. Log plot of the singular values is used to better separate the SV-spectra from each other.

3.2. Angle measure technique (AMT)

The angle measure technique (AMT) was introduced in 1994 [18] as a new method to characterise the complexity of geomorphic lines. AMT was designed to delineate changes in complexity of a geomorphic feature as a function of scale. We will use this approach by applying the AMT technique on an unfolded image to characterise periodic variations in the unfolded image pixel values. The AMT approach has profound implications for analysis of both 1-D and 2-D signal series containing a considerable amount of noise. AMT characterises the noise part as well as quasi-periodic phenomena of a measurement series in a novel fashion as a function of a scale-factor, s. For a more extensive explanation of the AMT used in chemometric modelling we refer to Ref. [11].

The AMT is implemented by the mean angle representation computed at a fixed scale and sampled at

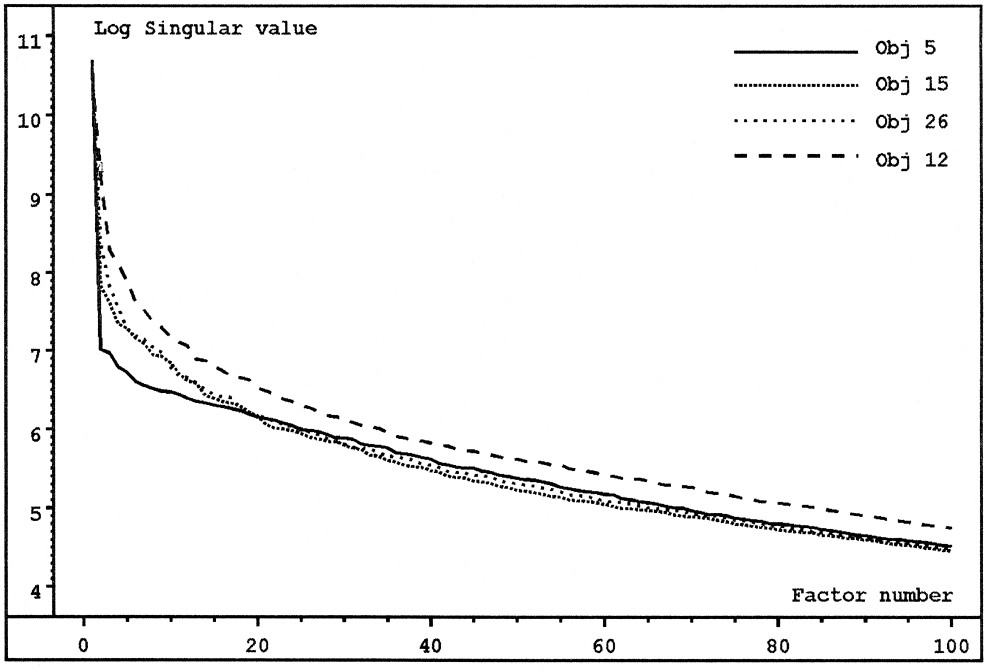

Fig. 3. The SV spectra of the four selected images of Fig. 1 are shown as a plot of the log (λ) against the singular value index. The SV spectrum of the low degree porosity is in the uniform random distributed region and the SV spectrum of the image with distinct holes is in the normal random distributed region.

a large number of points in the unfolded (vectorised) version of an image. (By an unfolding of an image **A**, we mean the vector obtained by aligning the pixel-rows of **A** sequentially after each other [19].) The AMT spectrum is in this manner a function of the scale-factor s. This is shown in Fig. 4.

Consider the image **A** of size $(m \times n)$. The corresponding vector is $\underline{\boldsymbol{A}}$ of length $l = mn$. The elements of $\underline{\boldsymbol{A}}$ range from $a_{\min}$ to $a_{\max}$ corresponding to the range of the pixel values in the original image **A**. Let P_i be a randomly chosen entry of the vector $\underline{\boldsymbol{A}}$. Let A_i and B_i be the points of intersection between the circle with radius s and the line constructed from the elements in $\underline{\boldsymbol{A}}$. Let $\boldsymbol{\alpha}_i$ be the supplement angle between the lines $A_i P_i$ and $P_i B_i$ calculated at point P_i. Then the AMT spectrum is defined by

$$\alpha(s) = 1/M \sum_{i=1,M} (\alpha_i)_s \tag{8}$$

where M is the number of random points sampled from $\underline{\boldsymbol{A}}$ and $(\alpha_i)_s$ is the angle calculated at point P_i with scale-factor s. It is often necessary to scale $\underline{\boldsymbol{A}}$ to obtain an optimal mean angle distribution. The normalised $\underline{\boldsymbol{A}}*$ is defined by

$$\underline{\boldsymbol{A}}^* = k(\underline{\boldsymbol{A}} - a_{\min})/|a_{\max} - a_{\min}| \tag{9}$$

where k is a scale factor to choose and $(\underline{\boldsymbol{A}} - a_{\min})$ means element-vice subtraction of the constant $a_{\min}$. The minimum and maximum pixel values of the matrix $\underline{\boldsymbol{A}}$ are given by $a_{\min}$ and $a_{\max}$. This normalisation makes $\underline{\boldsymbol{A}}^*$ scale independent and is often preferred before the raw vector $\underline{\boldsymbol{A}}$. The factor k will have to be determined from a pragmatic point of view.

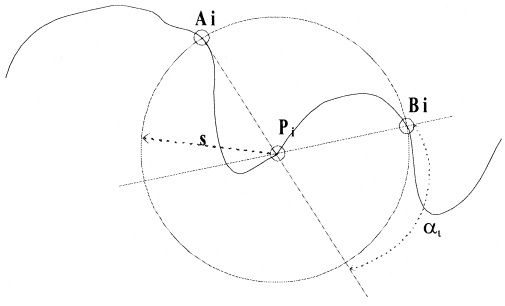

Fig. 4. The AMT spectrum is produced by the mean of several angle measurements along the curve as a function of the radius s in a set of random points (Esbensen et al., 1996 [11]).

Normally k is chosen in such a way that the AMT spectrum ranges from 0° to 100°. In some instances it is preferable to use only a scaled fraction, k, of the raw vector.

3.2.1. Implementation

An image **A** of size 256×256 pixels is vectorised into $\underline{\boldsymbol{A}}$ where $l = nm = 256 \times 256 = 65536$, $a_{\min} = 0$, $a_{\max} = 255$. The AMT was applied to each of the 96 images by selecting P_i, $(i = 1000)$ random points and calculating the mean angle by varying the scale s in steps from 1 to 256. The vector was multiplied by $k = 0.5$ to give a reasonable maximum mean angle range of 0–100. The randomisation was performed with different random seeds. Fig. 5 shows examples of AMT spectra for the four different images in Fig. 2. The unfolding technique used in this work will introduce a discontinuity at each 256th index point in the unfolded vector. In the AMT spectrum this will give a break point at every 128 scale value. Fig. 5 shows this situation. Systematic runs have shown that the visibility of the break point at every 128 scale value is less dominant as the constant k increases. The 'Y-space' scaling by selecting a proper value for the constant k is a part of an ongoing investigation of the effect(s) of different 'Y-space' scalings on the resulting AMT spectra. The discontinuity effect will be modelled as noise and will have no relevance to the reference (sensory porosity) [11].

3.3. The autocovariance (ACOV) and autocorrelation (ACOR) spectra

The Autocovariance and Autocorrelation spectra are attractive as multivariate features both because of their statistical interpretations as well as their potential of reflecting characteristic information of different textures. Texture-images containing large areas of small local variation (sample 26 of Fig. 2) generate slowly decaying spectra, while texture-images with globally dominant local variation (sample 5 of Fig. 2) will generate a faster decay. Fig. 6 shows the ACOV spectra for the four images of Fig. 2. Note that these spectra by definition include the magnitude of variation of the corresponding images. Fig. 7 shows the ACOR spectra of the same images, which are simply the 'scale-free' version of Fig. 6, suppressing the information on image-variance. The relationship

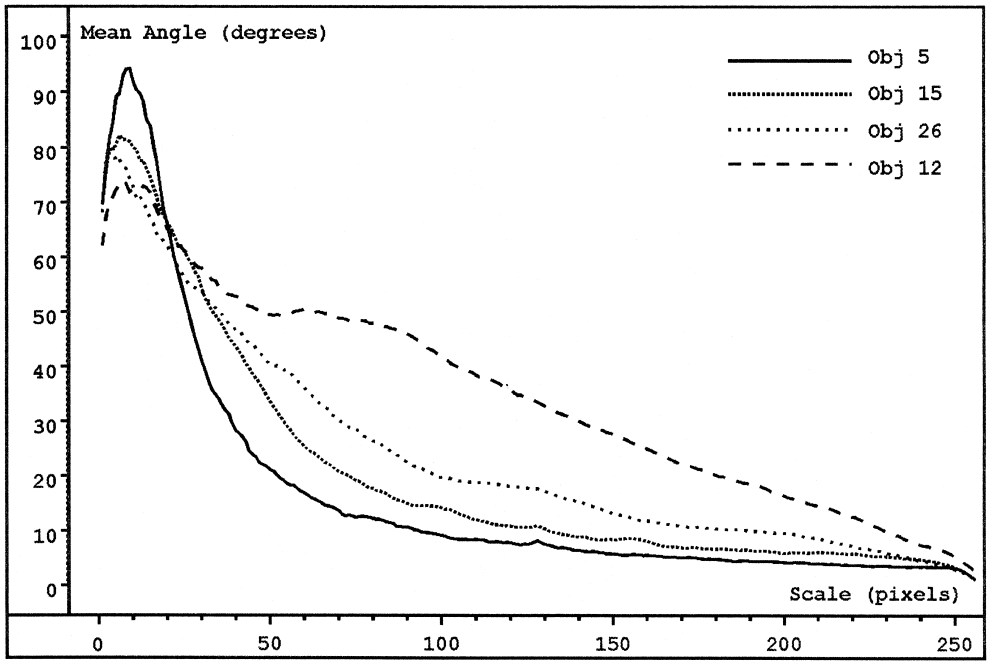

Fig. 5. AMT spectra estimated from the representative four images of Fig. 1. The degree of porosity is clearly distinct and visualised in the spectra.

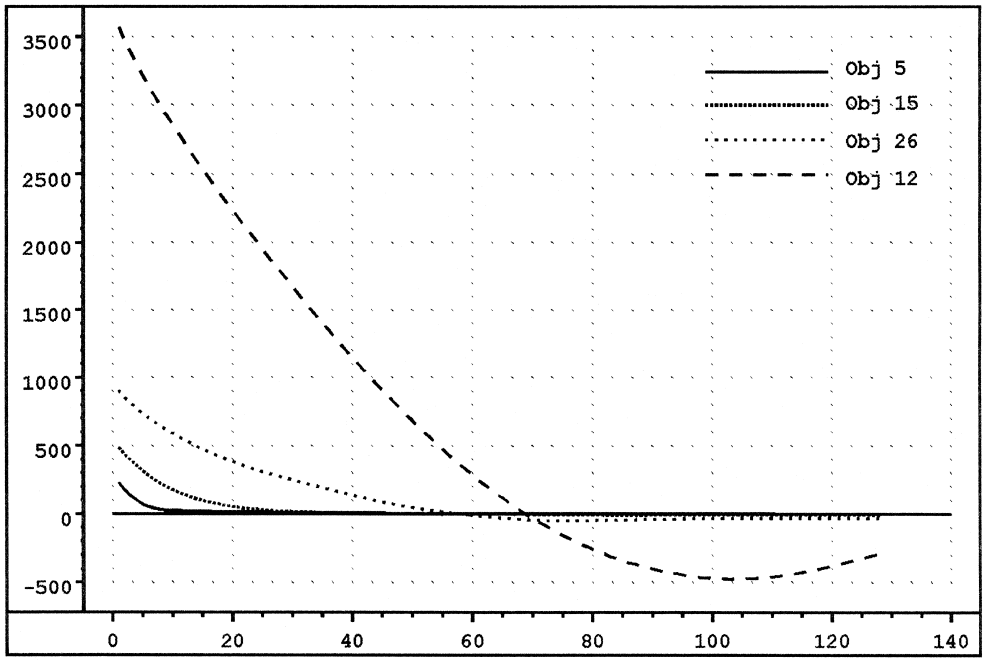

Fig. 6. Autocovariance spectra corresponding to the images of Fig. 1.

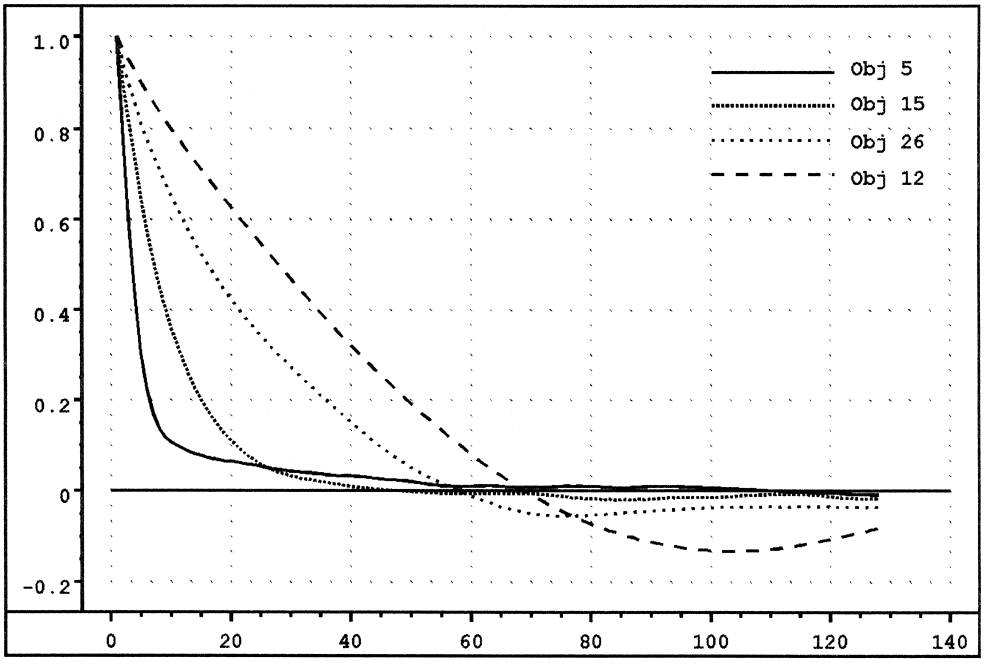

Fig. 7. Autocorrelation spectra corresponding to the images of Fig. 1.

$\mathrm{ACOV}(\mathbf{A}) = \mathrm{Var}(\mathbf{A}) \times \mathrm{ACOR}(\mathbf{A})$ express the formal correspondence between the two spectra. Here $\mathrm{Var}(\mathbf{A})$ is the magnitude of variation of an image $\boldsymbol{A}$. Which spectrum to prefer will depend on the response-variables dependency and sensitivity to the between-image variation.

The autocorrelation spectrum in terms of image analysis tradition [20] is defined as the empirical second moment of the noncentred and nonscaled image, and consequently different from its statistical analogue, but modulo an obvious linear transformation the two definitions are equal. By convention we will use here a statistical definition of the 2-D $\mathrm{ACOV}_{(m,n)}(\mathbf{A})$:

$$\mathrm{ACOV}_{(m,n)}(\mathbf{A}) \stackrel{\mathrm{def}}{=} 1/K \sum_{j,k} [\mathbf{A}(j,k) - \mu] \times [\mathbf{A}(j+m, k+n) - \mu], \quad (10)$$

where μ is the mean pixel value of $\mathbf{A}$ and (m,n) corresponds to a shift of the image by m rows and n columns and K equals the number of overlapping pixels between the shifted and the original image. As implicitly stated above, the statistical interpretation of this formula formally requires assumptions of stationarity (insensitivity to choice of origo). If isotropy (insensitivity to orientation) is also assumed, only the integer magnitude $r \approx \sqrt{(m^2+n^2)}$ of the shift (m,n) is relevant, and we can simplify the formula and obtain a 1-D version of the ACOV spectrum:

$$\mathrm{ACOV}_r(\mathbf{A}) \stackrel{\mathrm{def}}{=} 1/L_r \sum_{\{(m,n): r \approx \sqrt{(m^2+n^2)}\}} \times \mathrm{ACOV}_{(m,n)}(\mathbf{A}) \quad (11)$$

Here $r \approx x$ means the integer r closest to x and L_r equals the number of elements in the set $\{(m,n): r \approx \sqrt{(m^2+n^2)}\}$. Furthermore, the ACOR spectrum is defined as the ACOV spectrum divided by its first entry:

$$\mathrm{ACOR}_{(r)}(\mathbf{A}) \stackrel{\mathrm{def}}{=} \mathrm{ACOV}_r(\mathbf{A}) / \mathrm{ACOV}_0(\mathbf{A}). \quad (12)$$

(For completeness, note that $\mathrm{ACOV}_0(\mathbf{A}) = \mathrm{Var}(\mathbf{A})$.) Even in more general situations where assumptions of

isotropy and stationarity are invalid, both ACOV and ACOR can still be computed as direction-independent characteristic features.

The definition of $\mathrm{ACOV}_{(m,n)}(\mathbf{A})$ directly applied to large images leads to extensive computations. A mathematical argument using the Fourier transform and a computer-program executing the fast Fourier transform (FFT), however, offers a sufficiently fast estimation-procedure even in the case of image size 256×256. For a theoretical deduction of this result the reader is referred to Ref. [21], chapter 10. The authors present a 1-D argument that can easily be extended to the 2-D case required for images. A Matlab-routine implementing the FFT-based version for estimating $\mathrm{ACOV}_{(m,n)}(\mathbf{A})$ is given in Ref. [22].

Estimation of the ACOV and the ACOR spectra could also be based on a vectorisation of the images following the strategy of the AMT method. By a row-wise vectorisation we will implicitly focus exclusively on the horizontal structure and ignore the texture-information in other directions. Except for the possibly disturbing edge-effects of this strategy, the resulting spectrum should be identical to exclusively using the horizontal direction of the 2-D ACOV and ACOR spectra. Thus a vectorisation can very well be advantageous if the assumption of isotropy fails, but the texture of the images considered are arranged to have identical orientation (which indeed is the case for the data-set presented here).

3.3.1. Implementation

The images of size 256×256 pixels were loaded into MATLAB version 4.2. The fast implementation of Eq. (10) described in Ref. [22], resulted in 2-D ACOV spectra. Via Eq. (11) with $r = 128$, 1-D directional-independent spectra were derived. The ACOR spectra were obtained by division of each ACOV spectrum by its corresponding first entry.

3.4. Size and distances distributions (SDD)

A well-known method to present structural information from images of our kind is to extract texture primitives and quantify their characteristics, such as mean pixel intensity, size and shape [23,24]. Such a technique can be efficient if the images consist of well-defined textural elements which are easily segmented. In our case it is obvious that porosity is related to size and the number of pores/cm^2. This information can be extracted for each image as the distribution of pore sizes and the distribution of distances between all pores, represented as two histograms.

Using only these two histograms for multivariate calibration suffers from the limitation that they carry no information regarding the relation between the internal distances and their corresponding pore sizes. That is, the above extracted data do not reflect whether pores of different sizes are for instance evenly distributed or clustered in some fashion. One way to establish such information is to compute a relationship (for instance difference or ratio) between the pore sizes of each pair of pores and relate this to the corresponding distance between the pores. For each image we can then compute the average pore-size relation as a function of distance. By combining the resulting curve with the distance and size histograms we may have obtained a data set containing sufficient information for successful prediction of porosity.

3.4.1. Implementation

The images were imported into the software programme Image-Pro Plus as 8 bit greyscale bitmap files. The pores were mainly defined as the areas with pixel intensity below the threshold 155. However, the level of the mean pixel values of the grey level images may vary slightly due to different colours of the flour being used and also the treatment by the baking process. The threshold was therefore set manually for each of these images, guided by the mean pixel intensity, resulting in a threshold varying from 140–165. The lowest limit in size for what was counted as a pore was set to 1 pixel. After thresholding, a matrix was automatically generated for each image containing the size (number of pixels) and centroid coordinates for each pore. A program written for Matlab computed the distances between the pores and generated the histograms. Since the differences in pore size were very large, we generated the histogram of the natural logarithm of the sizes, ln(size). Figs. 8 and 9 show the two histograms generated from the four images in Fig. 1. The histogram of distances reaches high values when the number of pores in the image is large, as in sample 5. The histograms describing samples 12, 15 and 26 have differences in shape and

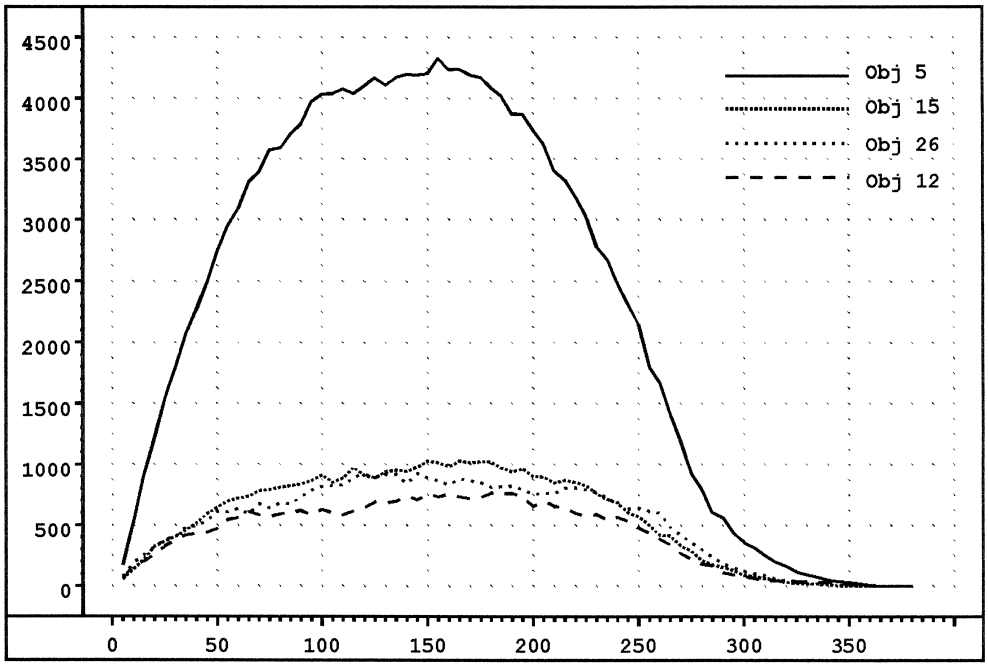

Fig. 8. Histograms of the distances between the pores of the corresponding images of Fig. 1. Bin size is set to 4.

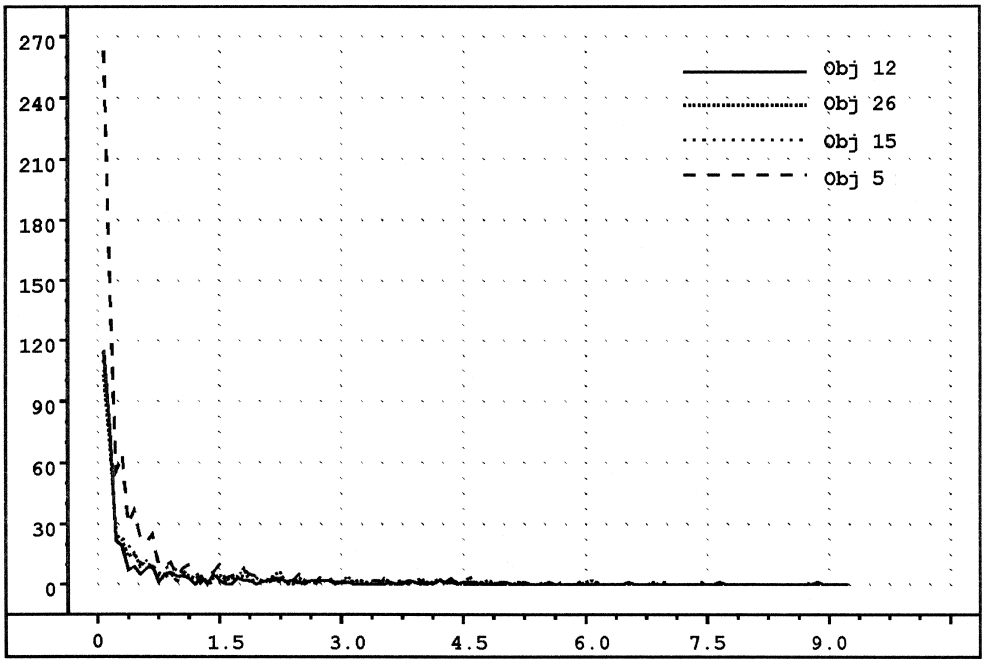

Fig. 9. Histograms of the ln(area) of the corresponding images of Fig. 1.

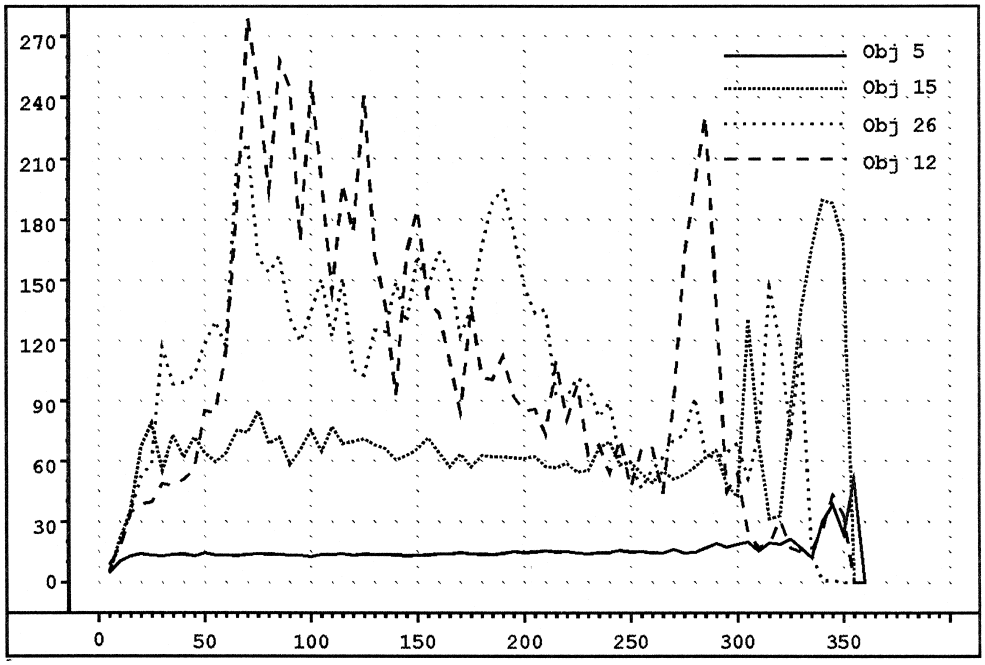

Fig. 10. Histogram of differences in pore size as a function of distance between the pores, $D(d)$, of the corresponding images of Fig. 1. Bin size is set to 5.

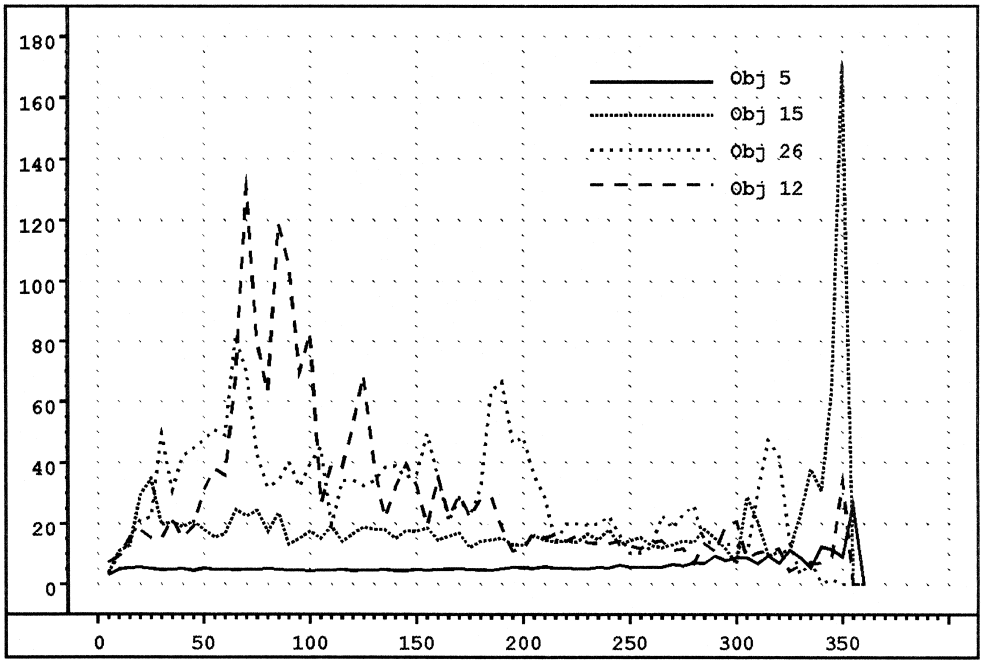

Fig. 11. Histogram of ratios of pore size as a function of distance between the pores, $R(d)$, of the corresponding images of Fig. 1. Bin size is set to 5.

level, but are small compared with sample 5. Fig. 8 shows that sample 5 contains a large amount of small pores. The very small responses above 6 correspond to the big pores in samples 12 and 26.

The relation between pore size as a function of distance was computed as follows: In a 256×256 pixel image the distance between two pores can range from 2–362. This range was divided into a number of intervals (histogram bins) of fixed size. For all distances within a given interval, the ratios and the absolute value of the differences between the pore sizes for the corresponding pore pairs were accumulated and divided by the number of distances within the interval. For each image, this created: (1) the average difference between pore sizes as a function of distance, $D(d)$, and (2) the average ratio between pore sizes as a function of pore distance, $R(d)$. Fig. 10 shows $D(d)$ and Fig. 11 shows $R(d)$ for the images in Fig. 1. The trends in $R(d)$ and $D(d)$ are quite similar. The histogram of sample 5 reflects small differences in pore size and that the pores are evenly distributed. The differences in pore size of sample 15 is bigger, giving a histogram of higher level, and the flat shape corresponds to an even distribution. Samples 12 and 26 both contain big and small pores resulting in large values of $R(d)$ and $D(d)$. The relatively great variations in these histograms are due to unevenly distributed pores.

4. Multivariate calibration

All the above feature extraction methods provide a data vector ($\boldsymbol{X}$) for each image. The focus in the present paper is to compare the ability of the different methods to predict external reference measurements of texture ($\boldsymbol{Y}$, in this case porosity). This is a typical multivariate calibration or regression problem. In mathematical terms, the problem can be stated the following way: Given measurements of both $\boldsymbol{X}$ and $\boldsymbol{Y}$ for a number of samples (in this case images), then build a relationship between them. The relationship can be used to predict $\boldsymbol{Y}$ from $\boldsymbol{X}$ and also to understand more about the relation.

In this paper we will only focus on linear relations, but more sophisticated methods based on, for instance, neural networks and locally linear models can also be envisaged. A linear regression model can be defined as follows:

$$y = b_0 + b_1 x_1 + \ldots + b_K x_K + e \qquad (13)$$

where the b's are regression coefficients to be estimated from the $\boldsymbol{X}$ and $\boldsymbol{Y}$ data for the calibration set of samples.

In the application to be presented below, the number of $\boldsymbol{X}$-variables is larger than the number of samples and the correlations among some of the different $\boldsymbol{X}$-variables are generally very high. This implies that standard least square regression can not be used for estimation of regression coefficients. The problem is called the collinearity problem [25].

Regression methods based on so-called data compression are often used to provide good relationships in the case of collinearity. They are shown to be well behaved in practice and can also be used to provide information about the relationship between $\boldsymbol{X}$ and $\boldsymbol{Y}$. In this paper we will focus on two of the most well-known methods among the data-compression methods, namely PCR and PLS. The PCR uses PCA for data and regresses $\boldsymbol{Y}$ on the principal components from $\boldsymbol{X}$. Data compression based on PLS uses $\boldsymbol{Y}$ as well in the estimation of components. This can sometimes lead to more parsimonious models, which can be useful for interpretation, but from a prediction point of view the two methods are usually quite similar. In the present paper main focus will be given to PCR, but some results from PLS will also be given.

We can then say that the data compression from the images goes through two steps, one for extracting features (data) from the images and one for extracting the main information from the feature vector.

The testing of the quality of the Eq. (13) will in all cases be performed by cross-validation. The validation structure is described at the end of Section 2.

The measure of the fitness of the model is given by R^2 which is the square of the square of the correlation of y and $\hat{y}$ (cross-validated) where y is the reference sensory porosity and $\hat{y}$ is the predicted porosity.

5. Results

The comparison between the different feature extraction methods is carried out without optimising the

single methods. We have chosen this automatic approach to compare the methods as independent as possible. It is often the practice to do data transformations according to a specific data structure to obtain better performance. In the PCR and PLS modelling there has not been any weighting transformations. Preprocessing like logarithmic transformations has not been done. In some instances such preprocessing is preferable and in our case might well lead to better performance for some of the methods. We have also investigated the potential of selecting different variable ranges. This is due to the fact that some methods need to focus on a subset of the total number of variables. All methods have the same range selections and thus not necessarily optimal for the each method. Ranges are selected from zero and to the number of variables in scope. No subranges were selected.

5.1. Results for the SVD

Table 1 shows the results of PCR and PLS modelling of porosity with the SVD representation of the images. By using the cross-validation method suggested the porosity is best explained after 6, 5 and 2 factors for the high, medium and low number of variables used. The porosity is best explained in PLS modelling at 5, 5, and 1 factors, respectively. The root-mean-square error of cross validation (RMSECV) is optimal at five components for both PCR and PLS using 100 variables. At 10 variables the RMSECV is much higher. The correlation factor reflects this tendency and gives 0.82 as correlation when all variables are in use. Sensory data will usually be somewhat noisy, so a correlation equal to 0.82 is a very high value. When compared with results obtained in a previous work, we found that there is a potential for an enhancement of the SVD method [4].

5.2. Results for AMT

The results of the AMT method are shown in Table 1. The results of PCR and PLS modelling of porosity are about equal when using high and medium number of variables while PCR is less precise than PLS when using a low number of variables. By using the cross-validation method suggested, the porosity is best explained after 7, 5 and 3 factors for the high, medium and low number of variables used using PCR. The porosity is best explained in PLS modelling at 4, 4, and 3 factors, respectively. The RMSECV is optimal at five components for both PCR and PLS using 100 variables. At 10 variables the RMSECV is significantly lower. The correlation fac-

Table 1
Results from the different methods compared

Method	Dataset	No. of variables	Opt. No. PC	RMSECV 3 segments	Correlation (R)
ACOR	ACOR50	50	2(2)	1.05(1.05)	0.78(0.78)
	ACOR30	30	2(2)	1.04(1.05)	0.78(0.78)
	ACOR10	10	5(5)	1.00(1.00)	0.80(0.80)
ACOV	ACOV50	50	5(4)	1.06(1.05)	0.77(0.78)
	ACOV30	30	4(4)	1.02(1.02)	0.79(0.79)
	ACOV10	10	6(4)	0.91(0.95)	0.84(0.82)
AMT	AMT256	256	7(4)	0.75(0.73)	0.90(0.89)
	AMT100	100	5(4)	0.77(0.76)	0.89(0.89)
	AMT10	10	3(3)	0.84(0.84)	0.86(0.88)
SDD	R(d) + DD	103	10(5)	0.80(0.79)	0.88(0.89)
	R(d) + DD	26	12(6)	0.87(0.88)	0.85(0.85)
	R(d) + DD	13	8(4)	0.93(0.97)	0.83(0.81)
SVD	SVD256	256	6(5)	0.97(0.96)	0.82(0.82)
	SVD100	100	5(5)	0.98(0.96)	0.81(0.82)
	SVD10	10	2(1)	1.25(1.24)	0.66(0.67)

The results are from systematic cross-validation in such a way that each replicate image block is used as a test set once. Results for PCR and PLS are shown.

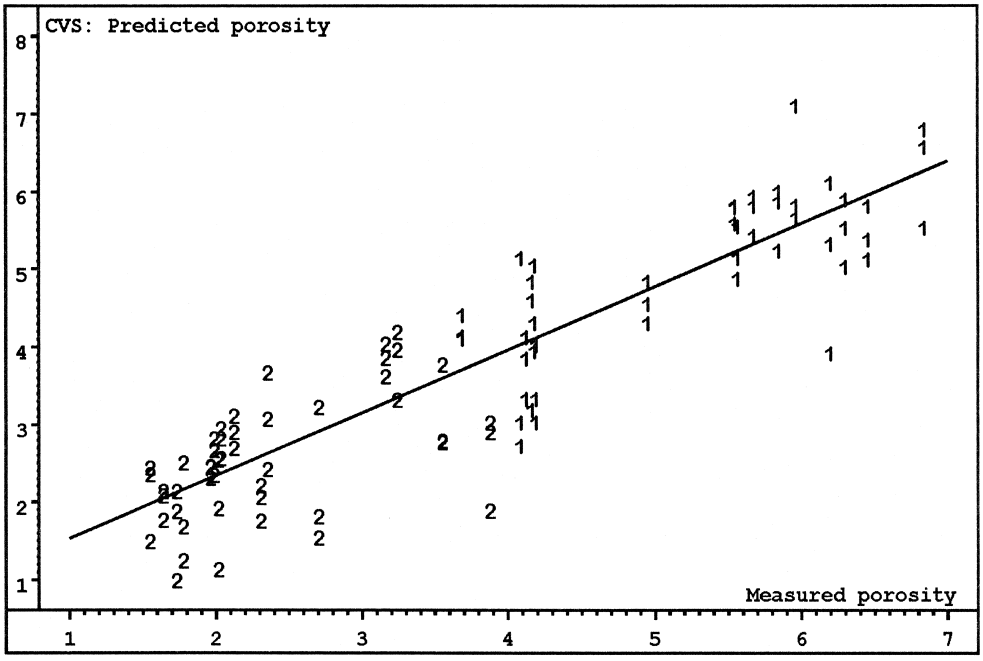

Fig. 12. The predicted porosity from the cross validation from the AMT method with all variables present. The object points are labelled with the corresponding baking method. We observe the classification of baking method.

tor reflects this tendency and gives 0.90 as correlation when all variables are in use. When compared to the SVD method this also indicates a very good modelling of the sensory data despite the noise level of such measurements. The predictions of porosity obtained from the AMT method are shown in Fig. 12. We observe that the feature extraction from images is capable to do a classification of both baking method and porosity.

5.3. Results for ACOV

Table 1 shows the results of PCR modelling of porosity with the autocovariance function representation of the images. Depending on the number of variables used from the spectra, porosity is best explained by 4–6 factors with correlations in the high 70s and low 80s of predicted vs. true porosity. The corresponding RMSECVs are found to be in the range 0.90–1.05. Using only the first 10 variables of the ACOV spectrum seems to give slightly better models.

5.4. Results for ACOR

The results of PCR modelling of porosity with the ACOR function representation of the images tends to be slightly poorer than for the ACOV function indicating the variance within each image to contain some extra information about porosity. For both 30 and 50 variables two factors give the best results, while for 10 variables five factors are preferred. Also the correlations for the ACOR are in the high 70s and low 80s of predicted vs. true porosity. The corresponding RMSECVs are found to be in the range 0.99–1.05. There is however a tendency of slightly poorer results compared to modelling with ACOV spectra. Also here using only the first 10 variables seems to give better models.

5.5. Results for SDD

The four histograms generated gave quite different results. The distribution of distances (DD) obtained a correlation of 0.67 with sensory porosity. The

distribution of sizes (DS) performed better with a correlation of 0.80 whereas the average difference and ratio as a function of distance ($D(d)$ and $R(d)$) performed best with correlations of 0.82 and 0.84 using seven factors. By combining the histograms in all possible ways, it turned out that a combination of DD and $R(d)$ resulted in the lowest RMSECV of 0.80 and a corresponding correlation of 0.88 with sensory porosity. As long as $R(d)$ or $D(d)$ was combined with DD, correlations up to 0.88 were obtained, combined with DS they were about 0.85.

The results listed above were achieved by using histograms with bin size = 1, i.e., DD, $D(d)$ and $R(d)$ were vectors of length 361. To reduce the data sets and processing time the bins in each histogram were increased. A bin size of 5 for $R(d)$ and $D(d)$ resulting in 71 variables each and a bin size of 12 for DD giving 32 variables, maintained the above results. Table 1 shows the results of the combination of $R(d)$ and DD. A further decrease in the number of variables led to slightly bigger prediction errors.

6. Discussion

We have shown that it is possible to extract features from images of a porous structure and model these to corresponding sensory attributes when using this design. The results should, however, be more general when our example is regarded as a population of all possible examples. The process of doing a straightforward modelling of the whole sample of images and predicting features of new unknown samples is very promising for on-line applications of the food industry, among several applications. This process is uncomplicated when compared to traditional texture derivation and analysis. The mean sensory porosity has been used to represent each sample's porosity and not the actual value of each sample. This will probably lead to better performance of the models, but this has no impact on the comparison of the modelling techniques.

The results listed in Table 1 state that to extract feature vectors from the images and then model these by multivariate methods is a powerful approach for extracting desired information from images. It must be pointed out that sensory data usually are quite noisy and that a correlation of about 0.9 is the maximum obtainable [26]. Therefore both AMT and SDD must be considered as successful methods. Fig. 13 shows the RMSECV for the methods in the test. Besides obtaining the lowest RMSECV, the AMT method needs only seven principal components to explain the variation in the sensory data. This indicates that the AMT is more robust and is also regarded as the best method for this application. Fig. 12 shows the predictions from the cross validations of sensory porosity for the best method, AMT. We also observe that it is possible to use this plot as a classification of the treatment (baking process) the bread samples are produced with. This verifies that it is possible to monitor more features at the same time in the modelling [4]. In on-line applications this is preferable.

The unfolding technique used to produce the AMT spectrum introduces a discontinuity at each 256th index point in the unfolded vector. This has clearly no effect as it will be modelled as noise and hence is of no relevance to the modelled sensory porosity. We conclude this from the fact that there is about the same correlation when using 256 variables as with 100 which is less than 128 where the effect is introduced in the AMT spectrum (Table 1).

It is, however, important to realise that the modelling goodness of any extracted spectrum will depend heavily upon the nature of the images subject to analysis. The choice of method must necessarily be based upon some heuristic considerations. Experience from a related paper [22] together with the results presented here indicate that while the AMT and SVD spectra seem superior for more irregular structures (typical for the texture of bread) the ACOR method perform well for images containing periodic phenomena (such as bulbs of fat in the microstructure of mayonnaise [22]).

The modelling of images by the use of feature vectors and further modelling with multivariate PCR/PLS techniques illustrates the different compression levels in the modelling process. When the features are extracted to give rise to a spectrum there will be an image level noise (ILN) to take into account. This is illustrated with the SVD method by the compression of images to produce the SVD feature spectra. The residual images described by the feature extraction will give an indication of the ILN and how

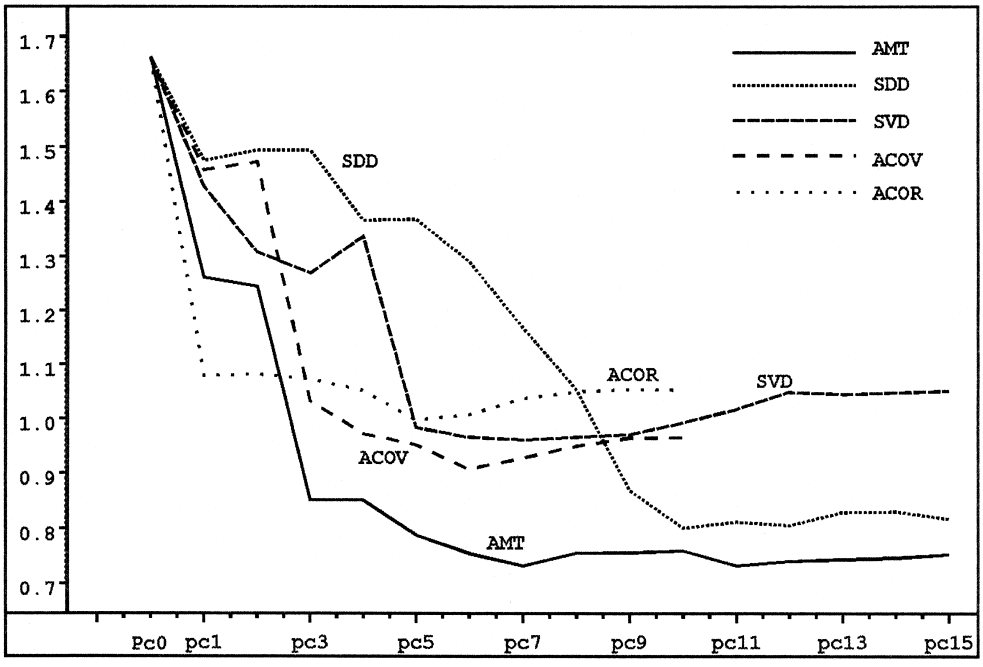

Fig. 13. The RMSECV representing the AMT, SVD, ACOR, AVOC and SDD methods at optimal performance. The AMT method performs best relative to the ACOR and SDD. The ACOR method was calculated with less factors than AMT and SDD.

many factors to take into account at this level. The optimal number of factors at this level is reflected in the different results of model ability in Table 1 [4]. When the next step of modelling is performed there will be a model level noise (MLN) which describes the residuals between the feature vector models and the true sensory porosity.

The process of choosing the variables to be used in the final modelling step is important to consider. The ILN should not be input to the PCR/PLS modelling of the sensory attributes. By reducing the ILN in front of modelling, it is possible to enhance the model performance. Of the image vectors we have used ACOV and ACOR seem to perform better when the last 40 variables are ignored, indicating that these variables are noisy. In SVD and AMT spectra, on the contrary, all variables in the full spectra seem to contribute more or less positively to the models since performance decreases with decreasing number of variables. In the SDD approach we reduced the variables by increasing the bin sizes which led to greater RMSECV. There may be a potential of improvement to keep the bin sizes narrow, detect possible noisy parts of the spectra and remove these from the histogram.

Each method has a possible potential of making a better performance in modelling the sensory porosity. We have compared the methods by not doing any optimisation or preprocessing especially designed for any of the feature extractions. In this way we have made the foundation of each calculation for each method as equal as possible. As mentioned and as the results illustrate, better results may be obtained by modelling subsets of the data, thus avoiding possible noise. Other improvements of the methods could be preprocessing of the images with special texture enhancing filters.

The SDD method is not a well explored extraction method. Other types of relationships rather than differences and ratios of pore sizes may be more appropriate to represent the desired information. Regarding other applications the choices of representation should be problem dependent.

The SDD spectra are quite easily interpreted when related to the pictures. Compared to the other methods the spectra can be more intuitively understood

which is an advantage for understanding the performance of the method.

All the feature vectors can be effectively computed from raw images. The SDD method as carried out here involved preprocessing of the images in which manual setting of the global threshold for each image was the most time-consuming and subjective part. There exists, however, a variety of thresholding techniques [27] which can be applied to automatically compute the optimal threshold (global or local) for each image guided by specific criteria. If the images suffer from varying levels of mean pixel intensities, they should be standardised by a well-defined method [20]. The whole procedure of generating the SDD histograms can then be quick and automatic. A more objective segmentation may also improve the performance of the method. We will also state that an extensive optimisation of the algorithm would be a violation in our wish to compare the methods as objective as possible.

The possible nonlinearity of the data structure has not been taken into account. It is possible to handle this situation by the use of nonlinear modifications to the modelling methods. It is also possible to achieve a better performance in modelling by the use of nonlinear neural networks. This has been shown to be a promising topic in sensory analysis and spectroscopic techniques. It is most likely to suggest that this is the case in feature modelling.

7. Conclusion

We have shown that it is possible to extract features from images of a porous structure and model these to corresponding sensory attributes. All images are modelled enabling the prediction of features on unknown sample images. This process is uncomplicated and more easily implemented when compared to traditional texture derivation and analysis methods. We have shown that it is possible to model sensory attributes like porosity. Other attributes may be modelled at the same time. Based on this we conclude that the AMT performs best in the modelling of sensory porosity. The fact that the data topology plays an important role in the choice of method, we conclude that typical porous structures visually judged is possible to model with all theses methods. The ACOR and ACOV methods performed poorly compared to the others, but as reported this method has a good potential when it comes to other data structures [22].

References

[1] M.R. Ellekjær, Rapid assessment of relevant quality parameters of meat products using near infrared spectroscopy, DSc Thesis 1993:20, Agric. Univ. of Norway (ISSN 0802-3220, ISBN 82-575-0199-9).
[2] T. Næs, E. Risvik (Eds.), Multivariate analysis of data in sensory science, Data Handling in Science and Technology, Vol. 16, Elsevier Science, the Netherlands, 1996.
[3] D.L.G. Bertrand, D. Marion, M. Devaux, P. Robert, Description of the textural appearance of bread crumb by video image analysis, Cereal Chem. 69 (3) (1992) 257–261.
[4] K. Kvaal, P. Baardseth, U. Indahl, T. Isaksson, Relationships between measurements and features extracted from images, in: T. Næs, E. Risvik (Eds.), Multivariate analysis of data in sensory science, Data Handling in Science and Technology, Vol. 16, Elsevier Science, the Netherlands, 1996, pp. 135–157.
[5] I.Y. Zayas, Digital image texture analysis for bread crumb grain evaluation, Cereal Foods World 38 (10) (1993) 760–766.
[6] R.M. Haralick, Statistical and structural approaches to texture, Proc. 4th Int. Joint Conf. Pattern Recognition, 1979, pp. 45–60.
[7] F. Van der Heijden, Image based measurement systems, Object Recognition and Parameter Estimation, Wiley, Chichester, 1994, ISBN 0-471-950629.
[8] B. Ashjari, Singular value decomposition texture measurement for image classification, PhD Thesis, University of Southern California, Department of Electrical Engineering, 1982.
[9] K.H. Esbensen, S. Wold, P. Geladi, Relationships between higher-order data array configurations and problem formulations in multivariate data analysis, J. Chemometr. 3 (1988) 33–48.
[10] K.H. Esbensen, P. Geladi, Strategies for multivariate image analysis (MIA), Chemometr. Intell. Lab. Syst. 7 (1989) 67–86.
[11] K.H. Esbensen, K. Kvaal, K. Hjelmen, The AMT approach in chemometrics—first forays, J. Chemometr. 10 (1996) 569–590.
[12] N. Cressie, Statistics for Spatial Data, Wiley, New York, 1991, ISBN 0-471-843369.
[13] P. Baardseth, K. Kvaal, P. Lea, E.M. Færgestad, M.R. Ellekjær, Flour quality and baking process affects baguette quality, Annual Meeting & Food Expo '97, Orlando, USA, 14–18 June 1997, Book of Abstracts, p. 75.
[14] R.C. Gonzalez, R.E. Woods. Digital Image Processing, Addison Wesley, 1992, pp. 196–197.
[15] MathWorks: MATLAB for Windows®, version 4.2 (1995).

[16] Camo: The Unscrambler® version 5.5, Software and Manual (1995).
[17] W.H. Press, Numerical Recipes in C, 2nd edn., Cambridge Univ. Press, London, 1992.
[18] R. Andrle, The angle measure technique: a new method for characterising the complexity of geomorphic lines, Math. Geol. 26 (1) (1994) 83–97.
[19] B.K. Alsberg, B.G. Remseth, Unfolding, J. Chemometr. 6 (1992) 135–150.
[20] W.K. Pratt, Digital Image Processing, Wiley, New York, 1991, pp. 557–598.
[21] P.J. Brockwell, R.A. Davis, Time Series and Methods, 2nd edn., Springer-Verlag, 1991.
[22] U. Indahl, T. Næs, Evaluation of alternative spectral feature extraction methods of textural images for multivariate modeling, J. Chemometr. (1998), in press.
[23] S. Wang, D. Velasco, A.Y. Wu, A. Rosenfeld, Relative effectiveness of selected texture primitive statistics for texture discrimination, IEEE Trans. Syst. Man Cybernet. SMC-11 5 (1981) 360–370.
[24] F. Tomita, Y. Shirai, S. Tsuji, Description of texture classification and segmentation, IEEE Trans. Pattern Anal. Mach. Intell. PAMI-4 2 (1982) 183–191.
[25] H. Martens, T. Næs, Multivariate Calibration, Wiley, Chichester, 1989, ISBN 0-471-909793.
[26] T. Næs, R. Solheim, Detection and interpretation of variation within and between assessors in sensory profiling, J. Sensory Stud. 6 (1991) 159–177.
[27] P.K. Sahoo, S. Soltani, A.K.C. Wong, Y.C.A. Chen, Survey of thresholding techniques, Comput. Vision Graphics Image Process. 41 (1988) 233–260.