Technical University of Denmark

DTU

Evaluation of Yogurt Microstructure Using Confocal Laser Scanning Microscopy and Image Analysis

Skytte, Jacob Lercke; Ghita, Ovidiu; Whelan, Paul F.; Andersen, Ulf; Møller, Flemming; Dahl, Anders Bjorholm; Larsen, Rasmus

Link back to DTU Orbit

DTU Library
Technical Information Center of Denmark

Evaluation of Yogurt Microstructure Using Confocal Laser Scanning Microscopy and Image Analysis

Jacob L. Skytte, Ovidiu Ghita, Paul F. Whelan, Ulf Andersen, Flemming Møller, Anders B. Dahl, and Rasmus Larsen

Abstract: The microstructure of protein networks in yogurts defines important physical properties of the yogurt and hereby partly its quality. Imaging this protein network using confocal scanning laser microscopy (CSLM) has shown good results, and CSLM has become a standard measuring technique for fermented dairy products. When studying such networks, hundreds of images can be obtained, and here image analysis methods are essential for using the images in statistical analysis. Previously, methods including gray level co-occurrence matrix analysis and fractal analysis have been used with success. However, a range of other image texture characterization methods exists. These methods describe an image by a frequency distribution of predefined image features (denoted textons). Our contribution is an investigation of the choice of image analysis methods by performing a comparative study of 7 major approaches to image texture description. Here, CSLM images from a yogurt fermentation study are investigated, where production factors including fat content, protein content, heat treatment, and incubation temperature are varied. The descriptors are evaluated through nearest neighbor classification, variance analysis, and cluster analysis. Our investigation suggests that the texton-based descriptors provide a fuller description of the images compared to gray-level co-occurrence matrix descriptors and fractal analysis, while still being as applicable and in some cases as easy to tune.

Keywords: image processing, microstructure, quantification, statistics, yogurt

Practical Application: Confocal laser scanning microscopy images can be used to provide information on the protein microstructure in yogurt products. For large numbers of microscopy images, subjective evaluation becomes a difficult or even impossible approach, if the images should be incorporated in any form of statistical analysis alongside other measuring modalities or sensory data. Instead, automated image texture analysis can be used to provide objective descriptions of the images, and we provide a comparative study for a broad range of the many image texture analysis available. All of the investigated techniques should be applicable for any type of pseudo homogeneous image structures.

Introduction

When considering fermented milk products, their microstructure is of great importance for the consumer experience (Muir and Hunter 1992). During milk fermentation, the milk protein (mainly casein) milk aggregates and forms a gel network in which water and fat are embedded. The microstructure of this network affects the general functionality of the gel as well as the textural properties (Lee and Lucey 2010).

Several process parameters, such as temperatures and ingredient composition, can significantly affect the microstructure of the final milk gel. Such effects can be observed as changes in appearance in microscopy images, which are often characterized by the amount of protein interconnectivity, pore size distribution, and tortuosity (Lucey and others 1998a).

Confocal scanning laser microscopy (CSLM) has become popular in the field of dairy science within the last 2 decades. It has been used for studying milk gels made from bacteria culture as well as glucono-δ-lactone (Lucey and Singh 1997; Lucey and others 1998a, 1998b). CSLM is well suited for studying the microstructure of milk gels because little sample preparation is needed, and using fluorescent probes, specific compounds, such as the protein gel network, can be targeted.

CSLM images of the protein gel network are often used to verify or interpret results from other measuring techniques such as rheology or sensory panels (Ozer and others 1999; Pereira and others 2003; Lee and Lucey 2004; Pereira and others 2006; Guggisberg and others 2007). Commonly, the appearance of the micrographs is analyzed subjectively, which becomes difficult or even impossible when large amounts of images are considered. However, combined with image analysis, CSLM images can be objectively quantified, and potentially used in statistical analysis on equal footing as other measuring techniques.

CSLM images of protein networks typically resemble image textures, which can be described as images containing periodic subpatterns and further appears homogenous at certain scales. It is therefore natural to choose image texture description techniques for characterizing the protein microstructure. It should be stated that image texture and the texture of a yogurt are 2 distinct

concepts. The image texture can be considered a way of interpreting the micrographs depicting the microstructure of the protein network, whereas the texture of a yogurt is related to the physical properties of the yogurt, and defined by the microstructure of the protein network (Lee and Lucey 2010).

A common image texture descriptor is the fractal dimension, which describes complexity of irregular objects as a ratio of the change in detail by the change in scale (Mandelbrot 1983). Previously, there has been much emphasis on fractal analysis in regard to food structure (Barrett and Peleg 1995). This also applies to milk gels, where fractal models have been suggested to mathematically model the kinetic behavior of the gel formation (Horne 1999). Fractal analysis has successfully been used to study the properties of range of different protein gels (Hagiwara and others 1998; Pugnaloni and others 2005; Dàvila and others 2007; Dàvila and Parés 2007; Kuhn and others 2010; Torres and others 2012). Nonetheless, Pugnaloni et al. (2005) recognized that although the fractal dimension can be a sensitive descriptor, it is not unique. Consequently, significant different structures can have same fractal dimension, making the validity of fractal analysis problem dependent. Thus, Pugnaloni et al. (2005), Dàvila et al. (2007), and Dàvila and Parés (2007) apply fractal analysis in conjunction with other morphological measures such as pore size distribution and lacunarity.

Varma and Garg (2007) state that although both fractal and multifractal analysis have been investigated for texture classification, they are often lagging behind the state-of-the-art descriptors primarily for 2 reasons. First to the nonuniqueness of the fractal dimension, and second that fractal dimension is often computed globally for an image. Multiple studies in material classification (Leung and Malik 2001; Hayman and others 2004; Varma and Zisserman 2005; Caputo and others 2005) suggest that such image textures are best described using statistical distributions of the so-called textons, which can be considered prototypical image features. For an image, all specific interest points, or features, are individually assigned to the most similar texton, and the final image descriptor is defined by the frequency distribution of texton occurrences.

With the aim of objectively quantifying the microstructure of milk gels, we compare a range of texton-based image descriptors relying on different design principles. These descriptors are compared to both common fractal analysis, a state-of-the-art fractal based image descriptor, and another widely applied technique, gray level co-occurrence matrix, which is commonly used for texture characterization in many different application areas including food science (Zheng and others 2006).

Our investigation is based on the microstructure of different yogurts made by different fat and protein content as well as different heat treatments and incubation temperatures—factors that are expected to influence the formation of the protein network. Our expectation is that a well-behaved image descriptor will enable a classification of the micrographs to the categories defined by our experiment.

We also aim at using the image descriptors to determine the significance of the experimental factors. Stand-alone fractal analysis can easily be carried over to an analysis of variance framework, as it outputs a single numerical value. However, this is not usually the case for the other investigated image descriptors, where the output is often high dimensional. If the number of observations is small compared to the vector dimensionality, it can potentially lead to a lack of degrees of freedom in a multivariate analysis of variance framework. Instead, we represent the image descriptors

Table 1–The 2^4 experimental design.

	Subset 1				Subset 2				Subset 3				Subset 4			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Fat content [1.5/3.5 g/100 g]	−	−	−	−	−	−	−	−	+	+	+	+	+	+	+	+
Protein content [3.4/4.4 g/100 g]	−	−	−	−	+	+	+	+	−	−	−	−	+	+	+	+
Pre-heat treat. [75/90°C/15 min]	−	−	+	+	−	−	+	+	−	−	+	+	−	−	+	+
Incubation temp. [39/43°C]	−	+	−	+	−	+	−	+	−	+	−	+	−	+	−	+

Minus and plus denote low and high factor levels, respectively. The actual factor level values are given in the left-most column. The entire design is replicated across 3 d. The subsets will be referred to in the section "Nearest Neighbor Classification."

by their mutual distances. Thus, we can transfer the image descriptors to a similar, albeit nonparametric framework (Anderson 2001). Also, high-dimensional data can be difficult to interpret in its original form, thus we utilize the distance representation to form a hierarchical clustering of the data, which is well suited for visual interpretation.

Materials and Methods

Experimental design

A triple replicated 2^4 experimental design was used to create the different milk gel formulas. The 4 factors spanned fat and protein content, and preheat treatment and incubation temperature. An overview of the experimental design and the factor levels is shown in Table 1. All formulas were prepared in 1-L batches, and the 3 replicates were created on 3 consecutive days. Systematic day-to-day variation was treated as a nuisance factor and was eliminated by incorporating the days as a block in the experimental design (Montgomery 2005).

The fat content was varied using homogenized ultra-heat treated (UHT) semi-skimmed and whole milk for the low-level and high-level, respectively. Protein content was changed by adding 3% (w/w) skimmed-milk powder (Lactalis Ingredients, Bourgbarré, France) in the high factor level. Preheat treatment was performed using an autoclave (Systec V-Series, Holm & Halby, Brøndby, Denmark). Immediately after preheat treatment the samples were put into a cold-water bath. The following day the samples were heated to their target incubation temperature in water baths (Lauda Ecoline E100, Lauda Dr. R. Wobser Gmbh & Co. Kg, Lauda-Königshofen, Germany), and hereafter inoculated with a yogurt bacterial culture (YO-MIX 863, DuPont Danisco Range, Copenhagen, Denmark). The pH development was measured every 5 min (CINAC pH Controller, Ysebaert Dairy Div., Frepillon, France). When a sample reached pH 4.6 it was stirred and stored at 5 °C in a plastic container for a week.

Confocal laser scanning microscopy

All micrographs were captured using a Leica DM IRE2 confocal scanning laser microscope (Leica Microsystems, Heidelberg, Germany). Before microscopy the yogurts were stirred gently. Protein was stained using a fluorescein-5-isothiocyanate (FITC) solution (2 g/L) in acetone, which was applied to a microscope slide. The slide rested until the acetone had evaporated, and yogurt was applied to the slide, and rested for 15 min at room temperature. Sample micrographs were captured using a 40× oil immersion objective (40× HCX PL Apo 1.25 Oil), FITC was excited at 488 nm using an Ar/Kr laser, and the emitted signal was collected

from 503 to 533 nm at a depth of 7 μm into the sample. To eliminate bias in the data acquisition, the CSLM images were randomly sampled; however, major artifacts such as air pockets were avoided. The grayscale images were captured at a resolution of 1024 × 1024 pixels, and each image covered an area of 375 × 375 μm. During recording each image line was captured 4 times and averaged to reduce noise. Ten images were recorded for each yogurt sample, making a total of 480 micrographs.

Examples of micrographs from each of the samples in Table 1 can be seen in Figure 1. These examples were chosen as the median image, from each sample, as determined by the best overall performing method in the section "Nearest Neighbor Classification." Light or gray pixels depict the protein network, whereas the dark pixels depict the pores. In general, the density and interconnectivity of the protein network reflects the strength of the formed gel network (Lee and Lucey 2004). Thus, sample 1 is expected to have a weaker network than sample 16, however for other samples it can be difficult to determine by the eye.

Image preprocessing

Image normalization. The pixel intensities in the micrographs correspond to the combined response from the local protein content, dye concentration, and focus. Thus, comparisons between pixel intensities across the entire data set may not be meaningful unless we apply image normalization. We normalized the pixel intensities of each image to have zero mean and unit variance—a process also known as *whitening*, which ensures invariance toward fluctuations in the mean intensity level and contrast of the image (Prince 2012).

Gel segmentation. A typical preprocessing step when assessing protein gel micrographs is to segment out the gel structures. In this paper the gel segmentation is used for the box counting method in the section "Fractal Dimension." However, the segmentation could also be used to extract morphological descriptors. We will not consider such descriptors in this paper.

As seen in Figure 1, the CSLM images contain 2 major region types: the protein gel network and the pores. Commonly, gel network and pores are segmented by thresholding the grayscale values. However, the grayscale values in CSLM images of protein gels are typically not of bimodal type (Mellema and others 2000; Pugnaloni and others 2005), which complicates the identification of a suitable threshold.

Looking at Figure 1, 3 major regions can be identified in the images: the internal pore structure, the protein gel, and the transitions between the 2 regions. We propose to use 3 convolutions that capture the uniqueness of the 3 different regions. A mean filter and a local entropy estimate were used to distinguishing between the intensity and uniformity of the gel and pore regions respectively,
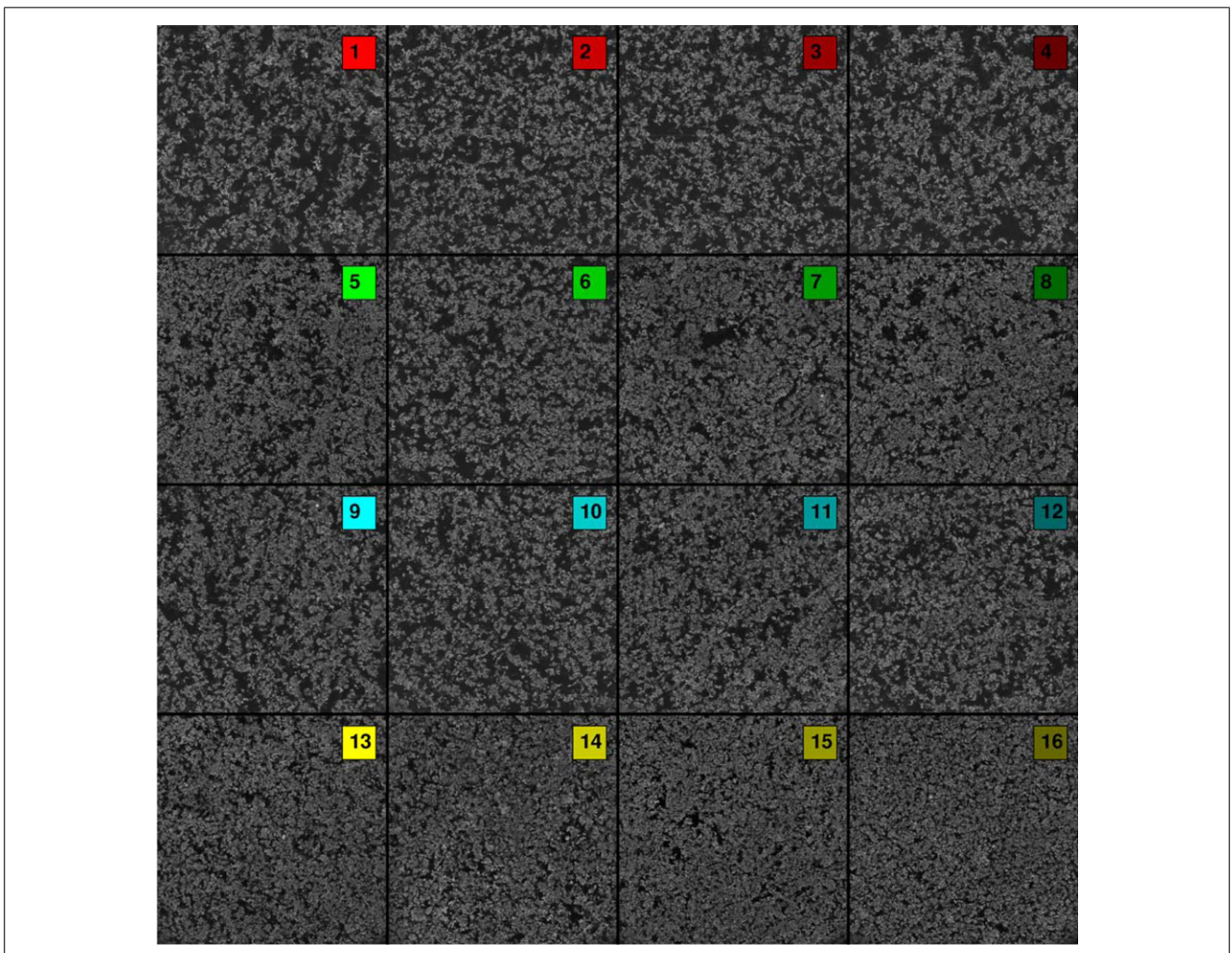

Figure 1–Examples of the 16 different yogurt samples from a single replicate. The numbering corresponds to that used in Table 1.

and a local skewness estimate was used to detect region transitions. To segment out the different regions, the 3-dimensional image representation was used as input for a Gaussian Mixture Model (Hastie and others 2009) with 3 clusters (one for each image region).

Image descriptors

A broad range of image descriptors are presented later. In sections "Intensity Histogram," "Fractal Dimension," and "Gray-Level Co-occurrence Matrices," we consider some classical image descriptors, and in sections "Local Binary Patterns," "Basic Image Features," "Image Patches," and "Local Fractal Features," we consider the image descriptors based on the texton approach, which is well suited for classifying image textures (Leung and Malik 2001; Hayman and others 2004; Varma and Zisserman 2005; Caputo and others 2005). Julesz (1981) initially proposed the texton approach, in which the main idea is to extract local descriptors at each pixel position in the image, and assign each descriptor into a set of predetermined descriptors (denoted textons). The final image descriptor is a normalized frequency histogram of the textons present in the image. In summary, textons can be seen as a dictionary of reoccurring image structures. Assigning each local descriptor to a texton can also be seen as a way of suppressing potential noise in the local descriptors.

For the methods in sections "Local Binary Patterns" and "Basic Image Features," the textons were defined *a priori* by the method, but for the methods in sections "Image Patches" and "Local Fractal Features," the textons had to be learned from the data set. This was carried out by randomly sampling local descriptors from the entire data set (denoted training data), followed by a *k*-means clustering on the training data. The *k* cluster centers hereafter comprised the textons. The texton pipeline is illustrated in Figure 2.

Looking at the microstructure in Figure 1, no distinct orientation of the structures can be seen. Thus, the image descriptor should preferably not be dependent on the direction of a given feature; that is, it should be rotational invariant. Furthermore, as suggested by Møller et al. (2013), the protein gels should also be considered at different scales for a more complete description. Table 2 gives an overview of the investigated image descriptors and their general features.

All methods mentioned in this section have one or more free parameters that can be chosen by the user. In most cases, these parameters significantly affects the performance of the method and must be chosen carefully. Specifically for the methods following the texton approach (without prior textons), the number of textons, found through *k*-means clustering, should be decided. Throughout the experiments we found that the performance of these methods was not overly sensitive to the number of textons.

Intensity histogram. This image descriptor is a histogram where each bin corresponds to a gray level in the image. The number of histogram bins can be varied by uniformly, downsampling the dynamic range of the gray levels.

Fractal dimension. Although there are many approaches to fractal analysis, most methods follows the same basic steps (Lopes and Betrouni 2009). First some image quantity, $N(r)$, is measured at different scales or step sizes, r. For true fractal scaling, these quantities can be related by a power law behavior:

$$N(r) \propto c \cdot r^D \quad (1)$$

where D is the fractal dimension and c is a constant. This expression is then linearized through a log-transformation:

$$\log(N(r)) = D\log(r) + \log(c) \quad (2)$$

and the fractal dimension can be estimated by the slope. For real world data, the log-transformed response may not show perfect fractal scaling (linear behavior) across all values of r. Thus, it is important to select the appropriate fractal range of r for estimating D. The intercept of Eq (2) is commonly referred to as *topothesy*, and describes the absolute magnitude of the pixel intensities (Russ 1994). Varma and Garg (2007) suggest including the intercept to alleviate the problem of the nonuniqueness of the fractal dimension. They denote it the *fractal length*, and we adopt this term for the rest of the paper. We will present results for the fractal dimension alone as well as with the fractal length included.

In this paper, we apply 2 different approaches to fractal analysis. First is the box counting method, which is the most popular and frequently used type of fractal analysis (Lopes and Betrouni 2009). Following the notations in Eq (1) and (2), boxes with side length r are placed on a binary image, and $N(r)$ corresponds to the number of boxes that sample the gel structure. In the 2nd approach, we look at the spatial autocorrelation modeled through semivariograms. For an image the semivariogram is estimated by

$$\gamma(h) = \frac{1}{2} E\left[(I(\mathbf{x}) - I(\mathbf{x} + h))^2\right] \quad (3)$$

where $I(\mathbf{x})$ is the pixel intensity at pixel coordinate $\mathbf{x}$ and h is a displacement. Again, we can relate back to power law behavior of Eq (1), such that $\gamma(h)$ and h correspond to $N(r)$ and r, respectively.
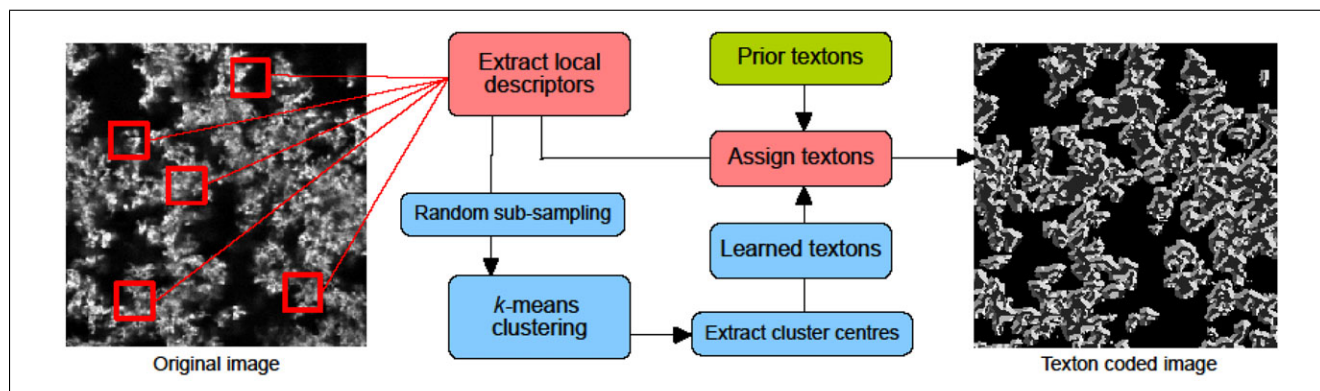
Figure 2–The applied pipeline for the texton approach. Local descriptors are extracted from the original images, and each descriptor is assigned a texton from a prior or a learned set of textons. The texton-coded image is made using 8 textons.

Table 2–Overview of the investigated image texture descriptors.

Descriptor name	Texton approach	Rotational invariant	Multiple scales
Intensity histograms	Yes (w. prior set)	Yes	No
Fractal (box count)	No	Yes	No
Fractal (variogram)	No	Yes	No
GLCM	No	Yes	Gaussian pyramid
LBP	Yes (w. prior set)	Yes	Gaussian pyramid
BIF	Yes (w. prior set)	Yes	Gaussian derivative filters
Image patches	Yes (w. training)	No	No
Local fractal features	Yes (w. training)	Yes	MR8 filter bank

GLCM, gray-level co-occurrence matrices; LBP, local binary pattern; BIF, basic image features.

A similar approach is used by Pugnaloni and others (2005). It should be clarified that we are estimating a fractal parameter for the image and not the actual fractal dimension of the protein gel itself.

Gray-level co-occurrence matrices. Gray-level co-occurrence matrix (GLCM) is a well-established method introduced by Haralick and others (1973). The GLCM expresses the probability of a given pixel intensity change between 2 pixels distanced by a displacement, h, and an angle, φ. Let (i, j) denote the 2 pixel intensities (or gray levels), then a GLCM, C is defined by $C = P((i, j) | h, \varphi)$. An example is given in Figure 3. Thereby, the number of gray-level intensities in the image defines the dimensionality of a GLCM. Typically, the number of grayscale intensities is downsampled. As in the section "Intensity Histogram," we used a uniform downsampling of the grayscale intensities, however other schemes can also be considered (Soh and Tsatsoulis 1999). The properties of the GLCM are commonly quantified by statistical measures, and Haralick and others (1973) proposes 14 different measures, which we also apply.

When creating the GLCM, multiple angles and displacements can be considered. Because of the lack of direction of the protein structures, we pooled the directional information by averaging the GLCM over multiple angles (0°, 45°, 90°, and 135°) as done in reference Soh and Tsatsoulis (1999). In terms of selecting appropriate displacements, Roberti de Siqueira and others (2013) shows it is viable to consider a single displacement at different scale representations of the image. One way to generate differ-

ent scale representations are Gaussian pyramids. Here, the original image is sequentially smoothed and downsampled (by a factor 2), which creates a series of different scales representations (or levels; Lindeberg 1994). Level 0 in the pyramid corresponds to the original image. Thus, the GLCM image descriptor was created using a single displacement at different levels in the Gaussian pyramid, and the final image descriptor was created as the joint distribution of the quantified GLCM across multiple scales.

Local binary patterns. Ojala et al. (2002) introduce the local binary patterns (LBP) for texture analysis. Here a circular neighborhood of radius, R, and P pixels is investigated for each pixel in the image. A local descriptor is defined by measuring the intensity difference between a pixel and its neighborhood. Negative differences are assigned 0, and positive differences are assigned 1. This results in a binary string (or pattern) of length equal to the amount of pixels in the neighborhood. Rotational invariance is obtained by shifting each binary pattern to get the minimum value. An example of the LBP descriptor is shown in Figure 4. Ojala and others (2002) notes that some pattern occurrences are generally more discriminative and they propose only considering a subset of so-called uniform patterns. The LBP image descriptor is a histogram over the occurrences of all binary patterns in the image. As for GLCM in "Gray-Level Co-occurrence Matrices," a similar multiscale approach using Gaussian pyramids was adopted (Qian and others 2011). Thus, the final image descriptor is the joint distribution of LBP histograms across multiple scales.

Basic image features. Crosier and Griffin (2008) introduce the basic image features (BIF) descriptor. This descriptor detects how local image structures change across scales. At first, 6 Gaussian derivative filters (from zero-order to 2nd-order) are applied to the image. From the filter responses each pixel is assigned 1 of 6 BIFs corresponding to bright/dark blobs, slopes, bright/dark lines, and saddle points. The BIFs are then assigned across multiple scales defined by the width of the Gaussian derivative filters. An example is shown in Figure 5. Crosier and Griffin (2008) find that 4 logarithmic distributed scales are appropriate. That is σ, 2σ, 4σ, and 8σ, where σ is the standard deviation of the Gaussian filters. The final image descriptor is constructed by considering how the
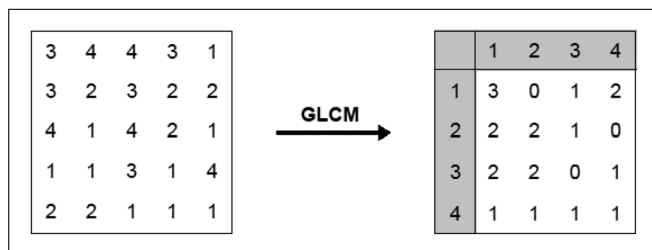
Figure 3–The gray-level co-occurrence matrix (GLCM) calculated for a 5 × 5 pixel image with 4 gray-level intensities. $h = 1$ and $\theta = 0°$.
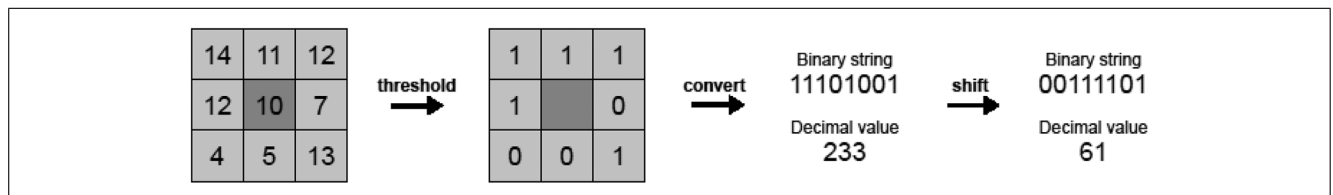
Figure 4–Local binary pattern (LBP) calculated for 3 × 3 pixel neighborhood ($R = 1$, $P = 8$). First the neighborhood is thresholded according to the center value, and hereafter the thresholded neighborhood is converted to a binary string. To obtain a rotation invariant description of the neighborhood, the binary string is shifted to obtain the smallest possible decimal value. (A) Original image (B) BIF (σ) (C) BIF (2σ) (D) BIF (4σ) (E) BIF (8σ).

BIF of each pixel changes across the 4 scales. This results in a frequency histogram with $6^4 = 1296$ bins.

Image Patches. Varma and Zisserman (2009) shows that the use of image patches as local features for the texton approach can be an efficient way of discriminating different materials. Image patches are extracted at each pixel location. From a patch of size $n \times n$ pixels, an n^2-dimensional local descriptor is defined as the joint distribution of pixels intensities. Each descriptor is contrast normalized by: $p \leftarrow p \left[\log(1 + L(p)/0.03) \right] / L(p)$, where p is the local descriptor and $L(p) = ||p||_2$ is the Euclidean norm of p, which has empirically been determined to give better discrimination. The texton approach with training is used to create the final image descriptor. An example of a texton dictionary based on image patches is shown in Figure 6.

Local fractal features. Rather than estimating the fractal dimension for the entire image, Varma and Garg (2007) propose to estimate the fractal dimension locally. Let $N(x, r) = \sum^I (\mathbf{x}, r)$, where $I(\mathbf{x}, r)$ are the pixel intensities within a closed disc of radius, r, at pixel coordinate, $\mathbf{x}$. Varma and Garg (2007) hypothesize that $N(\mathbf{x}, r)$ follows the power behavior of Eq (1), from which the fractal dimension can be estimated from the slope in Eq (2). Furthermore, they also use the fractal length (see section "Fractal Dimension") for the image descriptor. Also, as discussed in the section "Fractal Dimension," the fractal range, in which the fractal parameters are estimated, has to be determined. To accommodate rotation invariance and a multiscale representation, the MR8 filter bank (Varma and Zisserman 2005) is applied to the image. This results in 8 representations of the original image. For each representation, the fractal dimension and fractal length are estimated at each pixel location, thus a 16-dimensional local descriptor is obtained. As in the section "Image Patches," each descriptor was contrast normalized. The texton approach with training was used to create the final image descriptor.

Method evaluation

This section describes 3 ways of evaluating and interpreting the image descriptors presented in the section "Image Descriptors." The dimensionality of the image descriptor for each method and details about each descriptor will be provided in the section "Parameter Selection." A fair way of comparing descriptors of different dimensionality is to make comparisons based on the distances between image descriptors. When comparing histograms, a suitable distance measure is the χ^2 distance (Flannery and others 1992; Malik and others 1999):

$$\chi^2 \left(H^1, H^2 \right) = \sum_i \frac{\left(H_i^1 - H_i^2 \right)^2}{\left(H_i^1 + H_i^2 \right)} \tag{4}$$

where H denotes the histograms and i denotes the bin. Thus, it is a bin-to-bin comparison of the histograms, where pairwise difference between each bin is weighed by the numerical size of the bins. As a result, large and small bins are treated more equally. This distance measure is commonly used for texton frequency histograms (Malik and others 1999; Leung and Malik 2001; Varma and Zisserman 2005). We also found this distance measure suitable for the nonhistogram image descriptors used by fractal analysis and GLCM. Thus, all methods in this section evaluate and compare the image descriptors based on the χ^2 distance.

Nearest neighbor classification. The nearest neighborhood classifier is a popular evaluation method within the field of texture classification (Leung and Malik 2001; Cula and Dana 2004; Varma and Zisserman 2005, 2009). First the data are randomly split into training and test data equally for each class, and hereafter each observation in the test data is classified as the distance-wise closest observation in the training data. Multiple random splits of the data set are performed to give a statistical distribution of the classification rates.

Nonparametric multivariate analysis of variance. Within the analysis of variance (ANOVA) framework, multidimensional response vectors (image descriptors in this case) are typically handled using multivariate analysis of variance (MANOVA). As will be seen in the section "Parameter Selection," the dimensionality of the image descriptors will be as large as 1296. Considering the data set presented in the section "Experimental Design," this means that there are too few observations in order for the covariance matrices to be estimated in a feasible manner and an

A Original image   **B** BIF (σ)   **C** BIF (2σ)   **D** BIF (4σ)   **E** BIF (8σ)
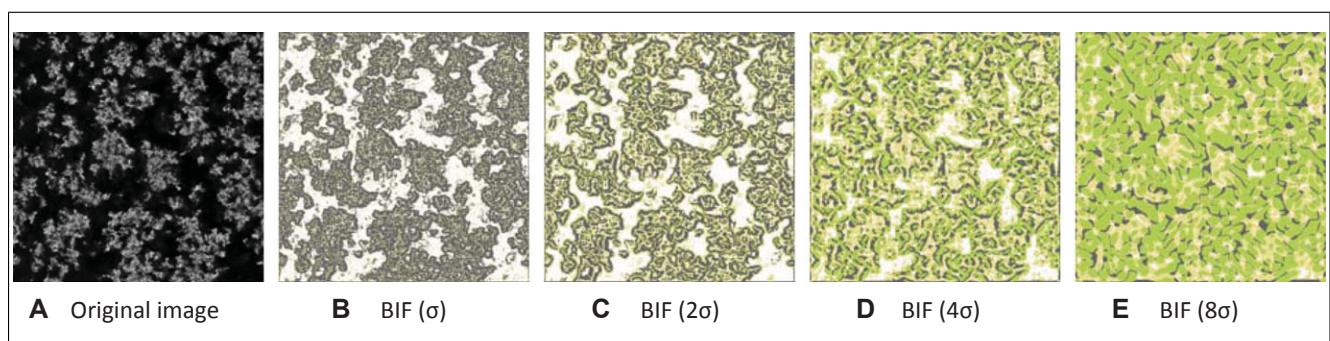
Figure 5–The basic image features (BIF) extracted at 4 different scales. The BIFs extracted from the original image are slopes (gray), bright blobs (white), bright lines (yellow), and saddle points (green). $\sigma = 1$ was used for the width of the Gaussian derivative filters.

Figure 6–An example of texton dictionary for 7 × 7 pixel image patches. The 64 textons are learned from the entire data set where 1000 random descriptors were extracted from each image. Note how certain patches appear as rotated/reflected version of each other, which is due to the lack of rotation invariance of the image patch method.

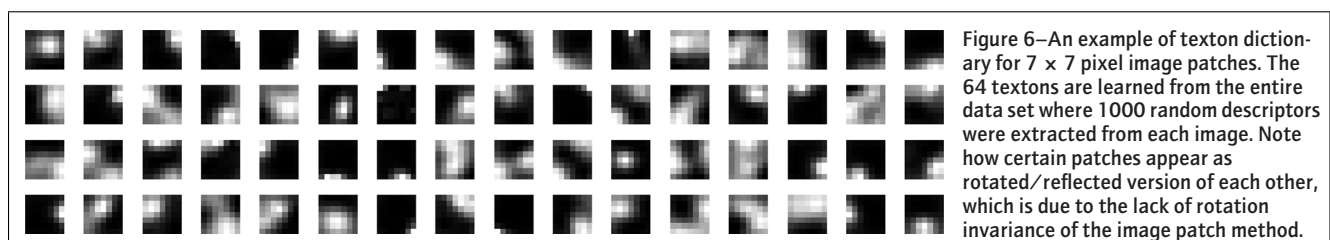alternative approach is needed. Such an alternative is proposed by Anderson (2001), and denoted nonparametric multivariate analysis of variance (NPMANOVA).

Here, the high dimensionality is handled by representing the multidimensional response vectors as a $N \times N$ distance matrix containing the pairwise mutual distances between observations. As a trade-off, this distance representation discards any information regarding the covariant structures in the data. From the distance matrix, the total variability can be partitioned by considering the mutual distances between observations rather than using the traditional sum of squares. A pseudo F-statistic can hereafter be calculated through permutation tests.

Contrary to ANOVA, NPMANOVA does not assume normality and is not necessarily based on Euclidean distances. However, similar to ANOVA, a major assumption for the NPMANOVA is equal variance across groups in the experimental design. This can be validated through a generalized version of Levene's test (Anderson 2006).

Hierarchical clustering. Hierarchical clustering is an unsupervised clustering method, which uses the interpoint distances to define a hierarchical ordering of the data. The clustering is typically done in a bottom-up fashion, meaning that each data point initially comprises a cluster and hereafter most similar clusters are greedily merged. Similarity is defined in terms of a link function (Hastie and others 2009). We used the group average link function in which similarity is determined by the average similarity between all members in the considered clusters. The final hierarchical structure can be presented as dendrograms.

Results and Discussion

Confocal scanning laser microscopy

Looking at the micrographs of the different samples in Figure 1 and the experimental factors in Table 1, several observations can be made. An increase in either protein or fat content generally increases the coverage of the protein network. The influence from fat is because of the homogenized fat globules being covered by a layer of protein (Yildiz 2010). This way they become an integral part of protein network. The influence of changing the protein content is straightforward.

Changes to preheat treatment and incubation temperature result in more subtle changes to the images. Higher preheat treatment temperatures result in more willing protein fusions, leading to more interconnected and dense gel networks (Lucey and others 1999). Incubation temperature affects the speed of the gel formation. Lower temperatures result in lower gel formation speeds, which again produce more interconnected and dense gels (Lee and Lucey 2004).

Gel segmentation

For the gel segmentation, we found filter sizes of 5×5 pixels to be appropriate, for the mean, skewness, and entropy filters. Segmentation results for 2 of the samples are shown in Figure 7. These segmentations successfully identify the protein, the pores, and the transitions between the 2 regions. In addition, the segmentations are found to be smooth. However, there seems to be a tendency of overestimating the transitions between pore and protein regions, resulting in transitions being favored more than the pores. This is most likely related to the depth of the pores. Some pores are deep and thereby represented by out-of-focus light, whereas some pores are less deep and creates light that is more in focus, thus creating structures similar to transitions. However, we found this to be of little importance, as we are only concerned with the segmentation of the protein, which was used for the box counting method in the section "Fractal Dimension."

Parameter selection

The free parameters for the methods in the section "Image Descriptors" were chosen such that each method achieved the highest correct classification rates, considering the entire data set, using the nearest neighbor classifier. However, tuning parameters using only certain parts of the data set could also be considered. To limit the parameter space substantially, the parameter search was sometimes limited to $[2^u, 2^v]$, where $u, v \in$. Table 3 summarizes the chosen parameters.

From Table 3, some interesting observations can be made. In general, most methods favor low spatial support, for example the optimal displacement for the semivariogram in fractal analysis was 6 pixels and the optimal patch size for image patch exemplars was 7×7 pixels. However, both LBP and GLCM successfully utilize multiple levels in the Gaussian pyramid, which can be seen as a way of increasing the spatial support. LBP consistently favored one or more levels in the Gaussian pyramid in addition to the original image. Also, although the chosen standard deviation for the basic image features seems small, the largest scale at 8σ corresponds to Gaussian filters with relative large support ($\approx 40 \times 40$ pixels). Thereby, GLCM, LBP, basic image features, and local fractal features successfully exploit the greater spatial support when describing the micrographs.

Looking at the texton-based methods, where the textons are learned through training, only a limited amount of textons are needed to provide a feasible description. This might be explained when considering the great similarity between the different protein gels. However, using the image patch textons requires twice as many textons compared to the local fractal features. This may be explained by the lack of rotation invariance of the image patches,

A Sample 3   **B** Sample 12   **C** Segmentation of Sample 3   **D** Segmentation of Sample 12
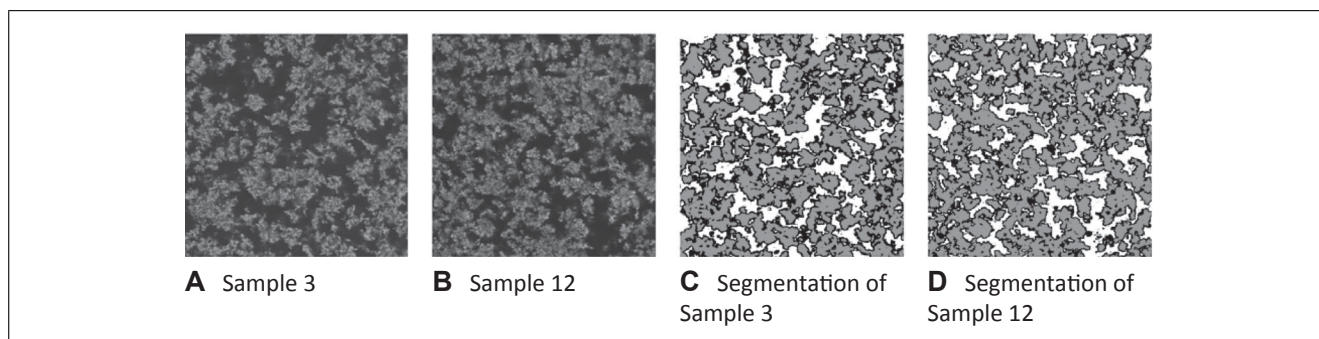
Figure 7–Segmentation examples. White denotes pores, gray denotes protein, and black denotes transitions between pores and protein. These examples correspond to close-up views of sample 3 and 12 from Figure 1.

Table 3–Summary of the major free parameters associated with each image descriptors.

Descriptor name	Parameter descriptions	Parameter values	Dimensionality
Intensity histograms	Number of greyscale intensities[a]	16	16
Fractal (box count)[b]	Fractal range, box side length (r)[a]	[1, 16] pixels	1
Fractal (box count)[c]	Fractal range, box side length (r)[a]	[1, 4] pixels	2
Fractal (variogram)[b]	Fractal range, variogram displacement (h)	[1, 4] pixels	1
Fractal (variogram)[c]	Fractal range, variogram displacement (h)	[1, 6] pixels	2
GLCM	Levels in the Gaussian pyramid	[0, 6]	154
	Number of greyscale intensities[a]	32	
	Displacement (h)	1 pixel	
LBP	Levels in the Gaussian pyramid	[0, 3]	144
	Binary pattern type	rotation invariant	
	Radius (R)	2 pixels	
	Number of sample points (P)	8	
BIF	Gaussian filter standard deviation (σ)	0.8	1296
Image patches	Patch size	7 × 7 pixels	128
	Number of textons[a]	128	
Local fractal features	Fractal range, disc radius (r)	[1, 5] pixels	64
	Number of textons[a]	64	

GLCM , gray-level co-occurrence matrices; LBP, local binary pattern; BIF, basic image features.
The rightmost column refers to the dimensionality of the final image descriptor.
[a]Denotes that the parameter search has been performed in a limited interval.
[b]Only the estimate of the fractal dimension is considered.
[c]Both estimates of fractal dimension and fractal length are considered.

Table 4–The average correct classification rates plus/minus one standard deviation for the different image descriptors.

Descriptor name	Entire data set	Subset 1 sample 1–4	Subset 2 sample 5–8	Subset 3 sample 9–12	Subset 4 sample 13–16	Superset subset 1–4
Intensity histogram	39.0 ± 10.8	44.3 ± 22.7	51.8 ± 22.4	50.6 ± 21.0	59.0 ± 22.1	74.5 ± 20.7
Fractal (box count)[a]	22.1 ± 09.2	28.5 ± 18.4	43.9 ± 23.1	50.5 ± 22.5	55.7 ± 21.6	54.0 ± 22.8
Fractal (box count)[b]	35.9 ± 10.4	38.0 ± 21.1	56.0 ± 22.6	51.5 ± 21.4	66.5 ± 20.8	68.8 ± 21.9
Fractal (variogram)[a]	36.4 ± 10.4	53.6 ± 20.6	76.5 ± 17.4	56.5 ± 20.0	73.1 ± 19.2	50.7 ± 24.6
Fractal (variogram)[b]	44.2 ± 10.7	57.5 ± 20.9	69.3 ± 20.4	65.5 ± 20.3	78.6 ± 18.2	62.5 ± 22.5
GLCM	48.8 ± 11.0	50.1 ± 22.4	58.4 ± 22.1	64.2 ± 21.1	76.6 ± 19.1	76.1 ± 19.8
LBP	58.1 ± 10.4	50.0 ± 21.5	71.5 ± 20.2	74.1 ± 20.2	89.8 ± 13.4	80.3 ± 18.6
BIF	69.5 ± 10.4	**61.6** ± 22.3	73.3 ± 19.9	72.8 ± 20.1	90.2 ± 14.4	**92.**0 ± 12.9
Image patches	**71.7** ± 10.3	60.9 ± 22.3	**76.7** ± 19.4	**82.5** ± 18.4	**94.0** ± 10.8	90.8 ± 13.9
Local fractal features	65.3 ± 10.6	60.4 ± 21.9	73.4 ± 20.1	76.9 ± 17.8	81.4 ± 18.2	88.5 ± 14.8

GLCM, gray-level co-occurrence matrices; LBP, local binary pattern; BIF, basic image features.
Classification rates are given for the entire data set, for subsets, and for a superset in which the subsets have been grouped. Values marked in bold indicate the highest average classification rate in each of the data sets.
[a]Only the estimate of the fractal dimension is considered.
[b]Both estimates of fractal dimension and fractal length are considered.

that is multiple patches are needed to describe multiple rotations of the same structure. The limited amount of unique image structures is further emphasized when looking at the methods covering intensity histograms or GLCM, which favors using only a few grayscale intensities.

Although we have considered and covered a broad range of these free parameters, it is by no means an exhaustive study of each method. For methods like GLCM and LBP, multiple extensions exist (Roberti de Siqueira and others 2013; Huang and others 2011). We have only considered a single extension of each method; the multiscale representation using Gaussian pyramids. However, several other multi scale or filter bank representations can be considered (Lindeberg 1994; Varma and Zisserman 2005).

Nearest neighbor classification

The results for the nearest neighbor classification are presented in Table 4. They were calculated based on 1000 random splits of the data, and averaged over the 3 blocks in the experimental design. The overall correct classification rates using all 16 samples are presented as well as smaller experiments using only subsets of the data set. Finally, classification rates are also given for an experiment where we created supersets of the previous subsets, that is sample 1 to 4 comprises a single class and so on. The subsets

were used to closely investigate the discrimination of the subtle changes made by preheat treatment and incubation temperature. The supersets were used to investigate discrimination between different fat and protein compositions.

Looking at the overall classification rates, the image patch method appears to be the best performer closely followed by the basic image features. In general, the texton methods appear to perform best. An important issue to note is the high standard deviations, immediately suggesting that the methods are not performing significantly different. However, the standard deviations are often comparable in size across the methods hinting that the variation, at least to some degree, can be explained by data variation. Referring to Figure 1, this can be further emphasized because of the potential overlap in appearance in the data set. The standard deviations are even higher when considering the subsets or grouped subsets, likely because of misclassifications having higher impact on each random split, when fewer groups are considered.

Looking at the performance on the subsets, it can be seen that the subtle changes to the gel, from the effects incubation temperature and preheat treatment, are easier to detect when the gel structure is dense. Subset 1 has the lowest correct classification rates, whereas subsets 2 and 3 have higher classification rates and

Table 5–The *P*-values for the NPMANOVA, the blocking of the 3 replicates is included as the factor *Day*.

Descriptor name	Day	Fat content	Protein content	Pre-heat treatment	Incubation temperature	Equal variance		
						Day 1	Day 2	Day 3
Intensity histogram	0.016*	<0.001*	<0.001*	<0.001*	<0.001*	0.137	0.078	0.08
Fractal (box count)[1]	0.513	<0.001*	<0.001*	<0.001*	<0.001*	0.349	0.021*	<0.001*
Fractal (box count)[2]	0.735	<0.001*	<0.001*	<0.001*	<0.001*	0.255	0.115	<0.001*
Fractal (variogram)[1]	<0.001*	<0.001*	<0.001*	<0.001*	<0.001*	0.484	0.266	0.029*
Fractal (variogram)[2]	0.005	<0.001*	<0.001*	<0.001*	<0.001*	0.381	0.233	0.009*
GLCM	0.010*	<0.001*	<0.001*	<0.001*	0.002*	0.359	0.249	0.16
LBP	0.044*	<0.001*	<0.001*	<0.001*	<0.001*	0.194	0.002*	0.161
BIF	0.040*	<0.001*	<0.001*	<0.001*	<0.001*	0.204	0.008*	0.054
Image patches	0.011*	<0.001*	<0.001*	<0.001*	<0.001*	0.18	0.114	0.071
Local fractal features	0.005*	<0.001*	<0.001*	<0.001*	<0.001*	0.146	0.144	0.089

GLCM, gray-level co-occurrence matrices; LBP, local binary pattern; BIF, basic image features.
The table also includes the tests for equal variance within each replicate.
*Denotes a significance on a 0.05 level.
[a]Only the estimate of the fractal dimension is considered.
[b]Both estimates of fractal dimension and fractal length are considered.

are generally comparable. The highest classification rates are found at subset 4 where the gel is most dense. Also, the superiority of the texton methods becomes more pronounced as the gel density increases.

The decrease in performance for sparser gels, as well as the large standard deviations, can also be explained by the fact that the textures in these gel structures are more inhomogeneous than the textures in denser gels (see images in Figure 1). This can be a problem as a single image may not contain a representative view of the actual sample, and 2 images of the same sample can potentially be very different. This could be mitigated using lower magnification during image acquisition or concatenation of several images.

Finally, when considering the superset, where each sample was classified into subset 1 through 4, we see large performance increases for all methods. This is especially noticeable in the classification results for intensity histograms, which see a large performance increase, as the amount of visible protein is the main changing factor across the subsets, which directly affects the grayscale intensity distribution.

In summary, the texton approaches seem to be well suited for discriminating between the microstructure of the different samples. However, because of the large standard deviations in classification rates, it is hard to determine the best texton method. Although the fractal analysis is outperformed by the texton

methods, it is still a strong competitor to the more complex GLCM approach, which additionally utilizes the multiple scales. Also, it can be seen that applying the fractal length in the fractal analysis increases the discriminative power for both fractal methods.

Nonparametric multivariate analysis of variance

Table 5 presents the results for the NPMANOVA. The blocking of the replicates was also included in the analysis as the factor *Day*. The test for equal variance among samples within each of the 3 replicates is also included. For both analyses, a square root transformation of the data was used to suppress some of the data variation.

Most image descriptors found that all factors in the experimental design were significantly changing the appearance of the protein network in the micrographs. Also, most image descriptors found significant day-to-day variation between the replicates, which can be expected when working with bacteria cultures. Looking at the test for equal variances, most methods showed significant differences in variance across factor groups within 1 or 2 of the replicates. This can be a problem as a difference in-group variance can lead to falsely detecting significant changes in-group means (Anderson 2001). Thus, the results from the NPMANOVA should be treated with some degree of caution, when considering the methods that show significant differences in variance between factor groups.
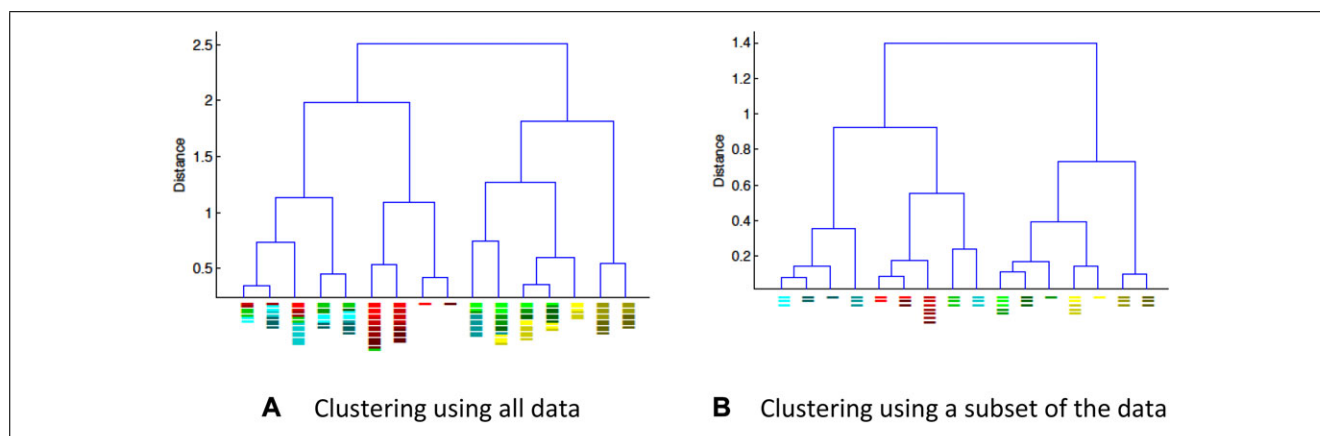

Figure 8–Hierarchical clusterings presented using dendrograms. The colors correspond to the sample colors in Figure 1, and each colored bar represents a sample image. In (a), the hierarchical clustering is performed on all data in a single replicate. In (b), the clustering is performed using only 3 selected observations from each class in a single replicate.

Hierarchical clustering

Hierarchical clustering is commonly presented using dendrograms. Figure 8 presents dendrograms from 2 different clusterings using the image patch descriptor from the section "Image Patches." For comparison, the clustering is based on the same replicate from which the example images in Figure 1 were extracted.

In Figure 8A, the hierarchical clustering is based on all data in this replicate. It can be seen that there is a lot of overlap between the different samples, which can be attributed to the potential overlap of sample appearance and within sample variation. It can be seen that subset 1 (red) is rather isolated from the other samples. There is a large overlap between subset 2 and subset 4 (yellow). This can be expected as both subset 2 and 4 had high protein content, which visually contributes most to the changes in the images in Figure 1. High fat content also changed the appearance but to a less degree than protein. This leads to subset 2 slightly overlapping with subset 3. By analyzing Figure 8 in greater detail, it can be seen that the densest sample in subset 2 (sample 6) favors subset 3, and the least dense sample in subset 3 (sample 11) favors subset 3.

Although Figure 8A gives an overview of the major changes in the data, the more subtle changes from different preheat treatments and incubation temperatures can be difficult to assess. To get a clearer view, some of image descriptors points can be filtered out. We did this by only selecting the median image descriptor and its 2 closest neighbors for each of the samples. The sample median was found as the image descriptor that was closest, in terms of Eq (4), to all other image descriptors within the same sample. The result is shown in Figure 8B, and comparing directly to Figure 1 there is a good visual correspondence between the images and the hierarchical representation.

Conclusion

When considering automated objective description of yogurt microstructure creating using different milk compositions and processing temperatures, the results of this paper suggest that it can be beneficial to consider texton-based image descriptors over conventional descriptors, such as fractal analysis and GLCM. Also, it was noted that although the conventional approaches performed similarly on the entire data set, the texton methods performed better when denser protein gel networks were considered. However, because of large within sample variation and overlap between samples, the best texton method could not be determined.

All of the applied image descriptors should be applicable for any type of microstructures, which resembles image texture, that is consists of pseudo homogeneous image structures when observed on a certain scale. When tuning the free parameters of the image descriptors, the majority of descriptors favored considering the protein microstructure on a small scale. However, certain methods also benefitted from observing the microstructure at coarser scales in addition to the small scale. Selecting optimal combinations of scale could potentially be investigated further.

We also showed 3 different ways of employing the high-dimensional image descriptor for classification. Nearest neighbor classification was used to evaluate the discriminative properties of the image descriptors and hierarchical clustering was used to give an interpretable view of the descriptors, which corresponded well to the visual impression of the micrographs. Finally, NPMANOVA was applied as a way of using the image descriptors in context of variance analysis. Here, most image descriptors agreed on the significant levels of the factors in the applied experimental design.

Although these methods rely on converting the high-dimensional image descriptors to distance-based representations, other approaches can be considered if the descriptors should be used in other forms of statistical analysis. Here, for example principal component analysis or multidimensional scaling could be employed to create low-dimensional representations of the image descriptors.

References

Anderson MJ. 2001. A new method for non-parametric multivariate analysis of variance. Austral Ecol 26(1):32–46.

Anderson MJ. 2006. Distance-based tests for homogeneity of multivariate dispersions. Biometrics 62(1):245–53.

Barrett AH, Peleg M. 1995. Applications of fractal analysis to food structure. LWT-Food Sci Technol 28(6):553–63.

Caputo B, Hayman E, Mallikarjuna P. 2005. Class-specific material categorisation. Proceedings of the IEEE Intl. Conference on Computer Vision, Beijing, China vol 2. p 1597–604.

Crosier M, Griffin LD. 2008. Texture classification with a dictionary of basic image features. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Anchorage, Alaska, USA. p 1–7.

Cula OG, Dana KJ. 2004. 3d texture recognition using bidirectional feature histograms. Intl J Comp Vis 59(1):33–60.

Dàvila E, Parés D. 2007. Structure of heat-induced plasma protein gels studied by fractal and lacunarity analysis. Food Hydrocolloids 21(2):147–53.

Dàvila E, Toldrà M, Saguer E, Carretero C, Parés D. 2007. Characterization of plasma protein gels by means of image analysis. LWT-Food Sci Technol 40(8):1321–9.

Flannery BP, Press WH, Teukolsky SA, Vetterling W. 1992. Numerical recipes in c. New York: Press Syndicate of the University of Cambridge.

Guggisberg D, Eberhard P, Albrecht B. 2007. Rheological characterization of set yoghurt produced with additives of native whey proteins. Intl Dairy J 17(11):1353–9.

Hagiwara T, Kumagai H, Nakamura K. 1998. Fractal analysis of aggregates in heat-induced bsa gels. Food Hydrocolloids 12(1):29–36.

Haralick RM, Shanmugam K, Dinstein IH. 1973. Textural features for image classification. IEEE Trans Syst Man Cybernet SMC-3 (6):610–21.

Hastie T, Tibshirani R, Friedman J. 2009. The elements of statistical learning, volume 2. Springer Science + Business Media, LLC, 233 Spring Street, New York, NY 10013, USA.

Hayman E, Caputo B, Fritz M, Eklundh JO. 2004. On the significance of real-world conditions for material classification. Proceedings of the European Conference on Computer Vision, Springer-Verlag, Prague, Czech Republic. p 253–66.

Horne D. 1999. Formation and structure of acidified milk gels. Intl Dairy J 9(3):261–8.

Huang D, Shan C, Ardabilian M, Wang Y, Chen L. 2011. Local binary patterns and its application to facial image analysis: a survey. IEEE Trans Syst Man Cybernet, Part C: Appl Rev 41(6):765–81.

Julesz B. 1981. Textons, the elements of texture perception, and their interactions. Nature 290:91–7.

Kuhn KR, Cavallieri ÂLF, Da Cunha RL. 2010. Cold-set whey protein gels induced by calcium or sodium salt addition. Intl J Food Sci Technol 45(2):348–57.

Lee W, Lucey J. 2004. Structure and physical properties of yogurt gels: Effect of inoculation rate and incubation temperature. J Dairy Sci 87(10):3153–64.

Lee W, Lucey J. 2010. Formation and physical properties of yogurt. Asian-Aust J Anim Sci 23(9):1127–36.

Leung T, Malik J. 2001. Representing and recognizing the visual appearance of materials using three-dimensional textons. Intl J Comp Vis 43(1):29–44.

Lindeberg T. 1994. Scale-space theory: A basic tool for analyzing structures at different scales. J Appl Stat 21(1–2):225–70.

Lopes R, Betrouni N. 2009. Fractal and multifractal analysis: a review. Med Image Anal 13(4):634–49.

Lucey J, Munro P, Singh H. 1999. Effects of heat treatment and whey protein addition on the rheological properties and structure of acid skim milk gels. Intl Dairy J 9(3):275–9.

Lucey J, Singh H. 1997. Formation and physical properties of acid milk gels: a review. Food Res Intl 30(7):529–42.

Lucey J, Tamehana M, Singh H, Munro P. 1998a. A comparison of the formation, rheological properties and microstructure of acid skim milk gels made with a bacterial culture or glucono-δ-lactone. Food Res Intl 31(2):147–55.

Lucey J, Teo CT, Munro PA, Singh H. 1998b. Microstructure, permeability and appearance of acid gels made from heated skim milk. Food Hydrocolloids 12(2):159–65.

Malik J, Belongie S, Shi J, Leung T. 1999. Textons, contours and regions: Cue integration in image segmentation. In: The Proceedings of the IEEE Intl. Conference on Computer Vision, Corfu, Greece. volume 2. p 918–25.

E: Food Engineering & Materials Science

Mandelbrot BB. 1983. The fractal geometry of nature. Henry Holt and Co. Henry Hold and Company, Inc., 175 Fifth Avenue, New York, NY 10010, USA.

Mellema M, Heesakkers J, Van Opheusden J, Van Vliet T. 2000. Structure and scaling behavior of aging rennet-induced casein gels examined by confocal microscopy and permeametry. Langmuir 16(17):6847–54.

Møller F, Larsen R, Carstensen JM. 2013. Imaging food quality [PhD thesis]. Technical Univ. of Denmark, Dept. of Informatics and Mathematical Modeling.

Montgomery DC. 2005. Design and analysis of experiments, 6th ed. John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA.

Muir DD, Hunter EA. 1992. Sensory evaluation of fermented milks: vocabulary development and the relations between sensory properties and composition and between acceptability and sensory properties. Intl J Dairy Technol 45(3):73–80.

Ojala T, Pietikainen M, Maenpaa T. 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 24(7):971–87.

Ozer BH, Stenning RA, Grandison AS, Robinson RK. 1999. Rheology and microstructure of labneh (concentrated yogurt). J Dairy Sci 82(4):682–9.

Pereira RB, Singh H, Munro PA, Luckman MS. 2003. Sensory and instrumental textural characteristics of acid milk gels. Intl Dairy J 13(8):655–67.

Pereira R, Merino-Matia L, Jones V, Singh H 2006. Influence of fat on the perceived texture of set acid milk gels: a sensory perspective. Food Hydrocolloids 20(2):305–13.

Prince SJD. 2012. Computer vision – models, learning, and Inference. Cambridge University Press, 32 Avenue of the Americas, New York, NY 10013-2473, USA.

Pugnaloni LA, Matia-Merino L, Dickinson E. 2005. Microstructure of acid-induced caseinate gels containing sucrose: quantification from confocal microscopy and image analysis. Colloids Surf B: Biointerfaces 42(3):211–7.

Qian X, Hua XS, Chen P, Ke L. 2011. Plbp: An effective local binary patterns texture descriptor with pyramid representation. Pattern Recog 44(10):2502–15.

Roberti de Siqueira F, Robson Schwartz W, Pedrini H. 2013. Multi-scale gray level co-occurrence matrices for texture description. Neurocomputing. 120:336–45.

Russ JC. 1994. Fractal surfaces. Plenum Press, New York. A Division of Plenum Publishing Corporation, 233 Spring Street, New York, NY 10013, USA.

Soh LK, Tsatsoulis C. 1999. Texture analysis of sar sea ice imagery using gray level co-occurrence matrices. IEEE Trans Geosci Remote Sens 37(2):780–95.

Soille P, Rivest JF. 1996. On the validity of fractal dimension measurements in image analysis. J Visual Commun Image Represent 7(3):217–29.

Torres IC, Amigo Rubio JM, Ipsen R. 2012. Using fractal image analysis to characterize microstructure of low-fat stirred yoghurt manufactured with microparticulated whey protein. J Food Eng 109(4):721–9.

Varma M, Garg R. 2007. Locally invariant fractal features for statistical texture classification. In IEEE Intl. Conference on Computer, Rio de Janeiro, Brazil Vision. p 1–8.

Varma M, Zisserman A. 2005. A statistical approach to texture classification from single images. Intl J Comp Vis 62(1–2):61–81.

Varma M, Zisserman A. 2009. A statistical approach to material classification using image patch exemplars. IEEE Trans Pattern Anal Mach Intell 31(11):2032–47.

Yildiz F. 2010. Development and manufacture of yogurt and other functional dairy products. CRC Press, Taylor & Francis Group, 6000 Broken Sound Parkway NW, Suite 300, Boca Raton, FL 33487-2742, USA.

Zheng C, Sun DW, Zheng L. 2006. Recent applications of image texture for evaluation of food qualities—a review. Trends Food Sci Technol 17(3):113–28.

E: Food Engineering & Materials Science