Discriminating Yogurt Microstructure Using Diffuse Reflectance Images

Jacob Skytte[1(✉)], Flemming Møller[2], Otto Abildgaard[1], Anders Dahl[1], and Rasmus Larsen[1]

[1] DTU Compute, Department of Applied Mathematics and Computer Science, Technical University of Denmark, Matematiktorvet B322, 2800 Kgs. Lyngby, Denmark
jlsk@dtu.dk
[2] DuPont Nutrition Biosciences ApS, Edwin Rahrs Vej 38, 8220 Aarhus, Denmark

Abstract. The protein microstructure of many dairy products is of great importance for the consumers' experience when eating the product. However, studies concerning discrimination between protein microstructures are limited. This paper presents preliminary results for discriminating different yogurt microstructures using hyperspectral (500-900nm) diffuse reflectance images (DRIs) – a technique potentially well suited for inline process control. Comparisons are made to quantified measures of the yogurt microstructure observed through confocal scanning laser microscopy (CSLM). The output signal from both modalities is evaluated on a 2^4 factorial design covering four common production parameters, which significantly change the chemistry and the microstructure of the yogurt. It is found that the DRIs can be as discriminative as the CSLM images in certain cases, however the performance is highly governed by the chemistry of the sample. Also, the DRIs shows better correlation to the CSLM images and are more discriminative when considering shorter wavelengths.

Keywords: Confocal scanning laser microscopy · Optical technique · Hyperspectral · Protein microstructure

1 Introduction

The quality perception and physical properties of fermented milk products are defined by the microstructure, formed through destabilisation and aggregation of the protein structures, during the milk fermentation process [1,2]. Confocal scanning laser microscopy (CSLM) is a popular microscopic technique for investigating the protein microstructure. This can be attributed the relatively easy sample preparation, which amounts to a fluorescent staining that is able to target the protein network directly, and produce highly detailed images of the protein microstucture [3].

However, while being practical, CSLM still requires sample extraction and preparation, which is not suitable for in-line process control during e.g. dairy

R.R. Paulsen and K.S. Pedersen (Eds.): SCIA 2015, LNCS 9127, pp. 187–198, 2015.
DOI: 10.1007/978-3-319-19665-7_16

production. In this regard, other optical techniques, which are non-invasive and non-contact, should be considered. Previously, transmittance and diffuse reflectance measurements have been used in great extent to monitor entire milk fermentations in relation to cheese making. Here, the initial structure formation is followed in order to predict the optimal cutting time, which results in a maximum cheese yield [4–6]. However, best to our knowledge little has been published in relation to discrimination of different protein microstructures. If possible, this potentially enables for detection of microstructural defects in dairy production, which can be beneficial for troubleshooting in the production line, and in general ensure a consistent product of high quality.

Thus, in the present work we seek to investigate the potential and limitations of using hyperspectral diffuse reflectance images (DRIs) for discriminating different yogurt microstructures. The DRIs are captured remotely, and are thereby non-contact, and requires no sample preparation – properties which are highly desirable when considering in-line applications. Previously, we have applied DRIs in relation to entire milk fermentations [7], and seen signals similar to those of transmittance and other diffuse reflectance techniques [5]. The discriminative properties of the DRIs in regard to microstructure, will be evaluated and discussed alongside CSLM images, quantified through image texture analysis. Both modalities are applied to a data set containing 16 unique yogurts created with different chemistry and microstructure, and has previously been used in [8]. Method evaluation is carried out by the means of classification and we furthermore attempts to correlate the signal output of the two modalities.

2 Materials and Methods

In this section, the experimental design is initially introduced, and hereafter the hyperspectral DRI technique is presented alongside the *decay parameter*, which is quantification of the DRIs into a single numerical value. Finally, a multidimensional CSLM image descriptor is briefly introduced alongside an approach to map this descriptor into a one-dimensional space, where it can be compared directly to the decay parameter.

2.1 Sample Design and Production

The samples were produced in a 2^4 factorial design spanning the four factors: *fat content, protein content, heat treatment temperature, and incubation temperature*. These factors are common process parameters that can be varied in yogurt production. The fat and protein content naturally affect the chemical composition of the yogurt, while the two latter factors define the amount of protein cross-linking that will be formed during the protein aggregation [3,9]. All factors were expected to affect the final microstructure. However, the heat-treatment and incubation temperature were only expected to result in subtle differences, as these factors only affect the amount of protein cross-linking.

The yogurt samples were produced by initially adjusting the fat- and protein content, and hereafter the milk was heat treated for 15 min at the target factor temperature level. After heat treatment, the milk was rapidly cooled in a cold-water bath. Finally, the milk was heated to the target incubation temperature, lactic acid bacteria were added to the milk, and the fermentation was initialised. The fermentation was continued until a pH of 4.6 was reached. Hereafter, the milk (now yogurt) was stirred and stored in a refrigerator for a week prior to measurements by DRI and CSLM. Three replicates of the experiment were performed on three consecutive days, resulting in a total of $3 \cdot 16 = 48$ yogurts. For each yogurt sample three DRI measurements were taken alongside 10 CSLM images.

The experimental design is summarised in Table 1, and Fig. 1 shows example CSLM images of the different microstructures for one of the replicates. For specific details on the experiment, and the data collection in relation to CSLM, please refer to [8].

Table 1. The 2^4 experimental design. 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 three days. The subsets, indicated in the first row, will be used during the method evaluation in Section 3.

	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/100g]	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+
Protein content [3.4/4.4 g/100g]	-	-	-	-	+	+	+	+	-	-	-	-	+	+	+	+
Heat treatment [75/90 °C/15min]	-	-	+	+	-	-	+	+	-	-	+	+	-	-	+	+
Incubation temp. [39/43 °C]	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+

2.2 Diffuse Reflectance Imaging

We have recently introduced a hyperspectral DRI system [7,10]. The system consists of a hyperspectral light delivery system made from a super continuum light source (*SuperK Extreme EXW-12, NKT Photonics, Birkerød, Denmark*) and an acousto-optic tuneable filter (*SuperK SELECT, NKT Photonics, Birkerød, Denmark*). The light delivery system produces a light beam (500-900 nm), which is focused on the sample surface, through a lens (focal length = 40 mm), at an oblique angle ($\approx 45°$). The spatial distribution of the resulting diffuse reflectance is captured using a CCD (*Grasshopper CCD, Point Grey Research, Richmond, Canada*). In front of the CCD a 6.5 cm spacer and an objective lens were installed (*23FM50L, Tamron Co. Ltd., Nagoya, Japan*), which produces an image of 1600×1200 pixels and a spatial resolution of 3.2 μm. A simplified schematic

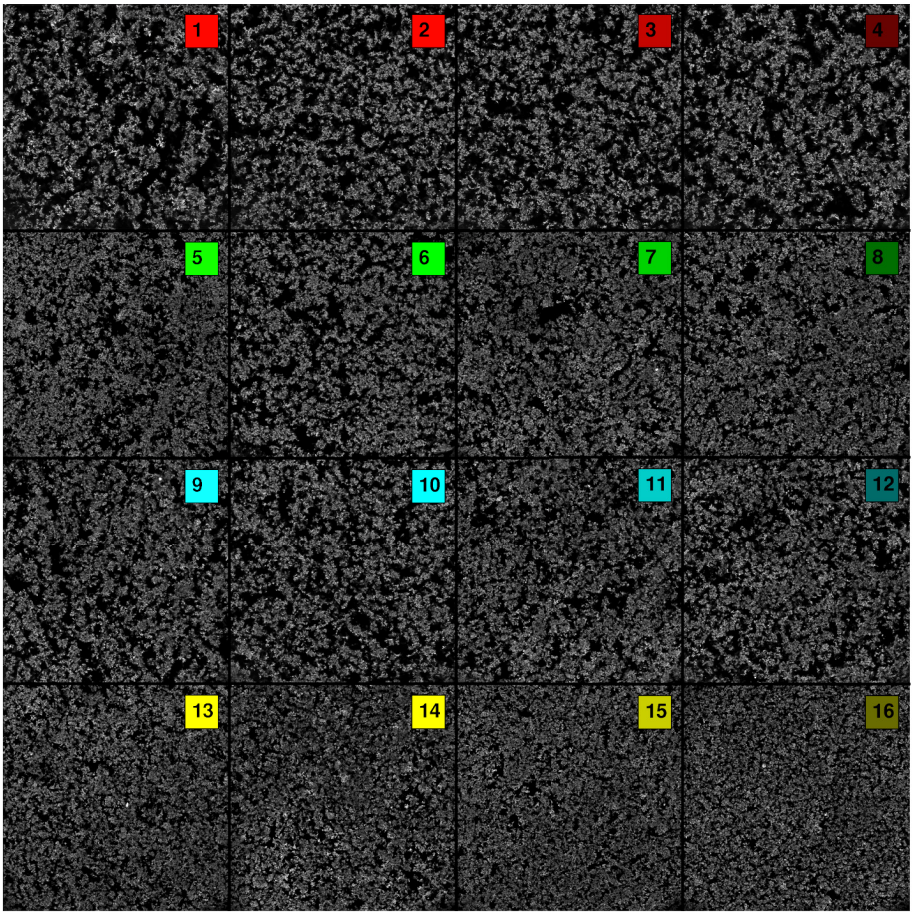


Fig. 1. Examples of the 16 different yogurt samples from a single replicate. The numbering corresponds to that of Table 1. Each image has a resolution of 1024×1024 pixels, and covers an area of 375×375 μm.

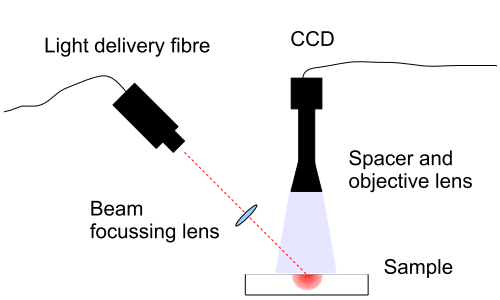
of the setup is shown in Fig. 2a, whereas Fig. 2b shows a captured image of semi-skimmed milk at 600 nm.

In order to quantify the DRIs we consider a light decay parameter [11], which quantifies the light intensity decay of the multiple scattered light, i.e. the light far away from the light incident point. The quantification scheme is illustrated in Fig. 2b and 2c. Here, the diffuse reflectance image is log-transformed twice and an intensity profile is extracted orthogonal to the direction of the incident light. A linear model is fitted to the outer part at both ends of the extracted profile, and the average slope of the two models (denoted the decay parameter from here on) describes the combined contribution of the absorption and scattering properties.

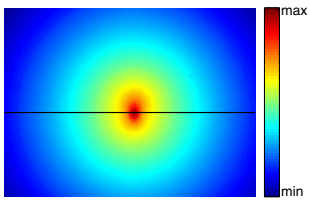
From light scattering theory it is well known that the main scatterer in milk and yogurt are the fat globules, due to their size [12,13]. However, the significantly smaller protein structures are also contributing to the overall scattering properties, and hereby the aggregation of the protein structures during fermentation can be observed by the scattering properties. This was verified in [14], where it was observed that the main optical difference between fermented and non-fermented milks is manifested in the scattering properties, while changes in the absorption properties are negligible. Thus, when considering dairy products, the decay parameter mainly reflects the scattering properties of the investigated sample.

2.3 CSLM Image Descriptor

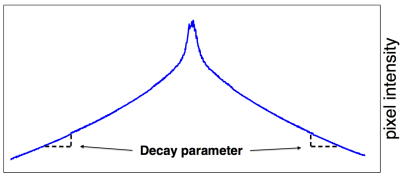
The discriminative properties for a broad range of image descriptors have previously been covered for CSLM images similar to those presented in Fig. 1 [8]. From said study, we select the overall best performing CSLM image descriptor, and compare the discriminative properties to those of the decay parameter. The image descriptor employs the *bag-of-words* approach [15]. Here, small image patches are sampled from the data, and a k-means clustering is performed to obtain a finite *visual vocabulary*, in which the cluster centres comprises the *visual words*. An image can hereafter be described by densely sampling image patches, and mapping them to the most similar visual word in the vocabulary.



(a) Schematic of system set-up



(b) Diffuse reflectance image



(c) Extraction of decay parameter

Fig. 2. (a) Illustrates the core components of the system set-up and (b) shows an example of a double logarithmically transformed diffuse reflectance image. In (c) an intensity profile of the diffuse reflectance image (corresponding to the black line in (b)) is shown, and the extracted decay parameter is visualised.

Thus, the image is described as a frequency histogram of common image structures in the data set. Following previous work [8], we apply an image patch size of 7×7 pixels and use 128 entries in the visual vocabulary. The distance between two image descriptors can be expressed using the χ^2-distance [16].

2.4 Multidimensional Scaling

The CSLM image descriptor described in the previous section is of high dimensionality. Therefore, in order to compare it directly to the decay parameter, the observations are mapped to a one-dimensional space. This can be done through multidimensional scaling, where the observations are considered in a distance space. Thus, rather than representing N image descriptors in their original p-dimensional space, the descriptors are represented as the mutual distances between all pairs of observations. Thus, the observations can be described in an $N \times N$ dissimilarity matrix, $\mathbf{D}$.

Multidimensional scaling can hereafter be used to determine a low-dimensional representation of the data, where the mutual distances of $\mathbf{D}$ are approximately retained. This can be done by seeking the values $\mathbf{Z} = (\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N)^T \in \mathbb{R}^{N \times k}$ (where $k < p$) that minimise the stress function:

$$S_{\mathrm{SM}}(\mathbf{Z}) = \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{(d_{ij} - ||\mathbf{z}_i - \mathbf{z}_j||)^2}{d_{ij}}, \tag{1}$$

where d_{ij} is an element of $\mathbf{D}$ and $|| \cdot ||$ is the Euclidean norm. The expression is a slightly modified version of the least squares formulation and denoted *Sammon mapping* [17], which puts more emphasis on preserving the small mutual distances. Sammon mapping was applied as some factors in the experimental design were only expected to make subtle changes to the microstructure.

2.5 Method Evaluation

To evaluate the discriminative properties of the decay parameter and the CSLM image descriptor, we apply the nearest-neighbour classification also used in [8]. This classification is performed on different partitions of the data set, in order to get an overview of the discriminative properties in relation to both chemical and microstructural composition of the yogurts. Additionally, we present the correlations between the one-dimensional representation of the CSLM image descriptor and the decay parameter extracted from the DRIs.

3 Results

3.1 Method Evaluation

Nearest Neighbour Classification Nearest neighbour classification was performed on the signal output from the decay parameter and CSLM image descriptor. For each class, a single observation was isolated (the test data), and

the nearest neighbour classification model was built on the remaining data (the training data). The presented test classification rates are based on the average over 1000 random splits of the data set, and further averaged across the three replicates of the experimental design. Additionally, the classification results were calculated for different partitions of the data set, in order to highlight the specific capabilities of the two modalities. the following partitions were considered:

- **Entire data set (sixteen classes)**. The overall performance on the entire data set was evaluated.
- **Superset (four classes)**. Here, each of the subsets, defined in Table 1, constitutes a single class, and reflects how well different chemical compositions can be discriminated.
- **Subsets 1-4 (each subset has four classes)**. The subsets defined in Table 1 were used to evaluate how well the subtle microstructural changes are reflected in the measurements, as the chemical composition of the yogurt changes.

Classification results for all partitions are summarised in Fig. 3 while Fig. 4 provides the corresponding confusion matrices for the CSLM image descriptor and the decay parameter at two different wavelengths.

Looking at the classification rates for the *entire data set* and the *superset* (Fig. 3), the CSLM image descriptor appears to perform significantly better than the decay parameter. When considering the *entire data set*, the classification rates for the decay parameter tend to decrease as the wavelength increases, and for the shortest wavelengths there is even a slight overlap in performance with the CSLM descriptor. This corresponds well to the presented confusion matrices (Fig. 4).

The same tendency is not seen for the classification rates of the *superset*. Here the performance generally appears to be lower, with larger standard deviations, and virtually no wavelength dependency. We found this was due to the samples with a high level of fat (samples 9 through 16), in which the effect of protein content seemed to be obscured. This can also be seen in the confusion matrices. This confounding, combined with only four classes, is likely to cause the lower performance.

Moving on to *subsets* 1 and 2, it can be seen that the performance actually appears similar for CSLM and the decay parameter, especially at the lower wavelengths. Considering the *subsets* 3 and 4, which have a high fat content, the performance is a bit lower for the decay parameter and a wavelength dependency is only seen for *subset* 4, which has a high protein content. The results for the *subsets* also seem to correspond well to the confusion matrices.

Correlation Before the decay parameter and the CSLM image descriptor could be compared directly, multidimensional scaling (see Section 2.4) was performed on the CSLM image descriptors. The average approximation error for mapping the CSLM image descriptor into a one-dimensional space was found to be around 11% when using the Sammon mapping and around 17% when using the least

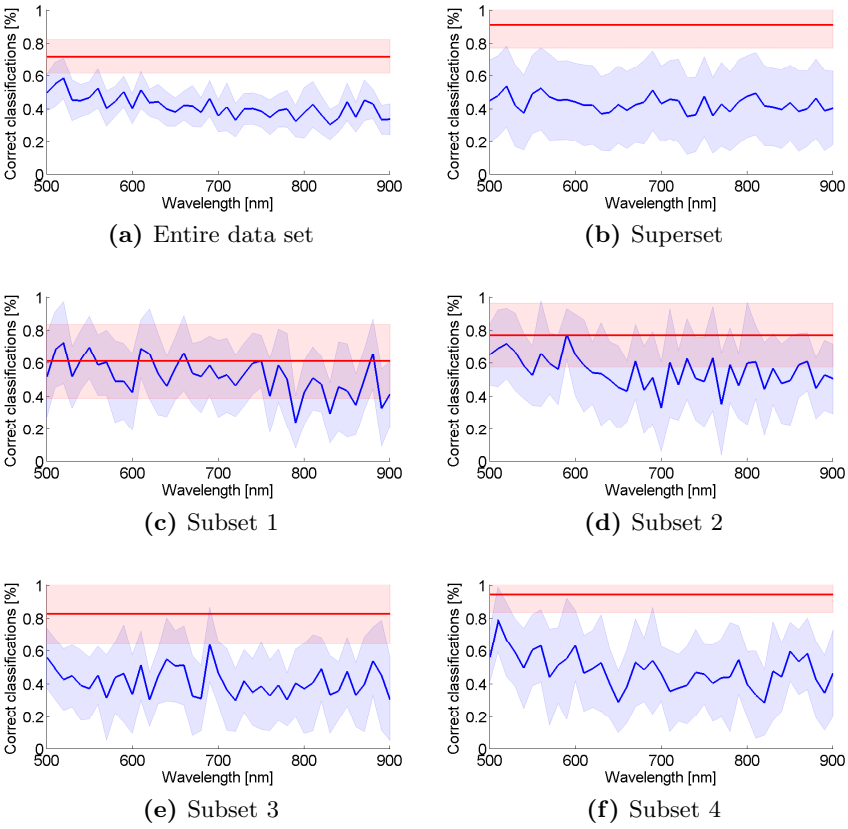


Fig. 3. Classification rates for the CSLM image descriptor (red) as reported in [8], and for the decay parameter (blue) for all considered wavelengths. The classification rates are provided for different partitions of the data set. The margins denote one standard deviation.

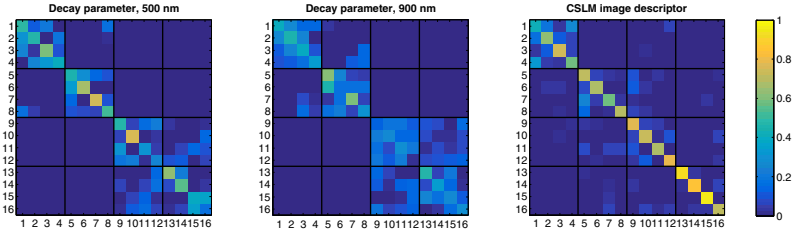
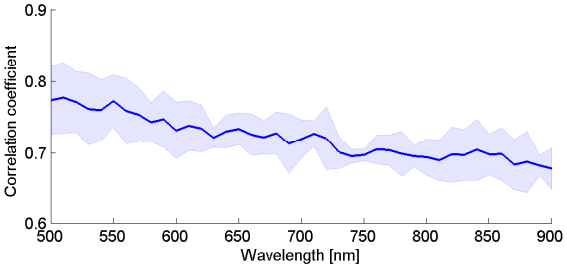


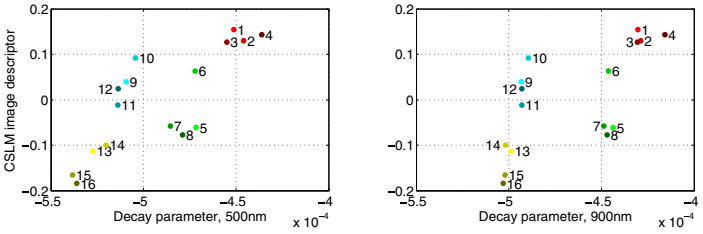
Fig. 4. Confusion matrices for the classification on the *entire data set*. The sample numbers corresponds to those of Table 1, and furthermore the grid highlights the four subsets defined in Table 1.

squares mapping. Hereby, the Sammon mapping was chosen for the multidimensional scaling. Hereafter, a correlation coefficient was calculated based on the average responses of the two modalities was, and further averages across the three replicates. The correlation coefficients are presented in Fig. 5a and again a wavelength dependency can be seen, and in this case it is even more prominent than for the classification rates.

The average responses, for a single replicate, are plotted in Fig. 5b against each other, at two different wavelengths. When comparing across wavelengths it quite clear how the signal at 500 nm is affected more by changes to protein content and the protein network microstructure compared to 900 nm. Especially for the samples with low fat content (sample 1 through 8) the effects of heat treatment and incubation temperature are pronounced. Contrary, at 900 nm there is a clear separation between low fat and high fat samples, while there is only a small effect from the protein content. Looking at the effects of heat treatment and incubation temperature they are more or less collapsed into a single point, making them hard to distinguish.



(a) Correlation coefficient across wavelength



(b) The two modalities plotted against each other

Fig. 5. (a) shows the average correlation between the average CSLM descriptor and the average decay parameter as a function of wavelength. The correlation is furthermore averaged across the three replicates and the margins denote one standard deviation. In (b) the average CSLM image descriptor and the average decay parameter plotted against each other at different wavelengths. The sample numbers and colours correspond to those of Table 1 and Fig. 1.

4 Discussion and Conclusion

We have assessed the discriminative properties of a single feature extracted from hyperspectral DRIs in regard to the protein microstructure of yogurts. Throughout the paper several observations can be made. The applied evaluation techniques generally showed a wavelength dependency, in which lower wavelengths correlated better to CSLM images of the microstructure and were better at discriminating between the different yogurt microstructures. However, when additional fat was added to the system, the microstructural effects from changing the protein content, heat treatment and incubation temperature diminished.

These findings correspond well to Mie theory [18] as the scattering efficiency of milk fat is known to be significantly larger than that of milk protein due to the difference in structure size; fat lies within the range 100-10000 nm while protein lies within 20-400 nm [12]. Additionally, the scattering contribution becomes smaller as the wavelength increases, when small structures (relative to the wavelength) are considered (Rayleigh scattering). For the applied wavelength interval (500-900 nm) this means that the scattering contribution from the fat structure remains fairly constant across the interval, whereas the contribution from the protein structures decays significantly across the interval. Thus, the effect of protein content and different protein microstructures should be more visible at lower wavelengths. This corresponds well to the observations of the evaluation techniques.

Comparing the classification performance between the two modalities, CSLM appeared to be more discriminative, however the classification performance was comparable in certain cases. This was especially clear for the samples 1 to 4 with low fat and protein content. Here, it should be noted that the CSLM have been captured at a single magnification level, and from Fig. 1 it can be seen that the less dense microstructure also appears more irregular. This suggests that the applied zoom level does not necessarily yield a representative view of the less dense microstructures and other zoom levels should be considered [8]. A similar, albeit inverted, problem was observed in relation to the decay parameter, where a high fat content seemingly confounded the scattering contribution from the other experimental factors. Here, a higher spatial resolution might be able to capture the appropriate dynamics in the depicted light diffusion. These observations highlight that the two applied modalities observe the samples at different scales.

Optical sensors have commonly been investigated in relation to the initial structure formation during cheese making. Claesson and Nitschmann [19] initially recommended near-infrared (NIR) light, as higher wavelengths showed a larger relative increase in the signal output, compared to lower wavelengths, during the structure formation. In the study by O'Callaghan et al. [20] commercially available sensors, for following the initial structure formation, are compared, and all investigated optical sensors are based on NIR light. In this regard the observations of this paper are remarkable, as they suggest that lower wavelengths should be favoured when the task is to discriminate between different yogurt microstructures. Thus, for this particular task, the results of this paper suggests

that new optical sensors should be developed, rather than relying on the commercially available sensors used in cheese making.

In conclusion, we believe the results of this paper encourages further investigation of using short-wavelength DRIs as means of inline process control for detecting microstructural artefacts in dairy production, especially when low-fat dairy products are considered.

References

1. Muir, D.D., Hunter, E.A.: Sensory evaluation of fermented milks: vocabulary development and the relations between sensory properties and composition and between acceptability and sensory properties. International Journal of Dairy Technology **45**(3), 73–80 (1992)
2. Folkenberg, D., Dejmek, P., Skriver, A., Ipsen, R.: Relation between sensory texture properties and exopolysaccharide distribution in set and in stirred yoghurts produced with different starter cultures. Journal of Texture Studies **36**(2), 174–189 (2005)
3. Lee, W., Lucey, J.: Formation and physical properties of yogurt. Asian-Australasian Journal of Animal Sciences **23**(9), 1127–1136 (2010)
4. Lucey, J.: Formation and physical properties of milk protein gels. Journal of Dairy Science **85**(2), 281–294 (2002)
5. O'Callaghan, D.J., O'Donnell, C., Payne, F.: Review of systems for monitoring curd setting during cheesemaking. International Journal of Dairy Technology **55**(2), 65–74 (2002)
6. Castillo, M.: Cutting time prediction methods in cheese making. In: Encyclopedia of Agricultural, Food, and Biological Engineering, pp. 1–7 (2006)
7. Skytte, J.L., Nielsen, O.H.A., Andersen, U., Møller, F., Carstensen, J.M., Dahl, A.B., Larsen, R.: Monitoring optical changes during milk acidification using hyperspectral diffuse reflectance images. In revision (2014)
8. Skytte, J.L., Ghita, O., Whelan, P.F., Andersen, U., ller, F.M., Dahl, A.B., Larsen, R.: Evaluation of yogurt microstructure using confocal laser scanning microscopy and image analysis. Journal of Food Science (in press, 2015)
9. Lucey, J., Munro, P., Singh, H.: Effects of heat treatment and whey protein addition on the rheological properties and structure of acid skim milk gels. International Dairy Journal **9**(3), 275–279 (1999)
10. Nielsen, O.H.A., Dahl, A.L., Larsen, R., Møller, F., Nielsen, F.D., Thomsen, C.L., Aanæs, H., Carstensen, J.M.: Supercontinuum light sources for hyperspectral subsurface laser scattering. In: Heyden, A., Kahl, F. (eds.) SCIA 2011. LNCS, vol. 6688, pp. 327–337. Springer, Heidelberg (2011)
11. Carstensen, J.M., Møller, F., Frisvad, J.L.: Online monitoring of food processes using subsurface laser scattering. In: Advances in Process Analytics and Control Technologies, pp. 5–7 (2009)
12. Walstra, P., Walstra, P., Wouters, J.T., Geurts, T.J.: Dairy science and technology, 2nd edn. CRC Press (2010)
13. Martelli, F., Del Bianco, S., Ismaelli, A., Zaccanti, G.: Light propagation through biological tissue and other diffusive media: theory, solutions, and software. SPIE Press (2010)

14. Nielsen, O.H.A., Subash, A.A., Nielsen, F.D., Dahl, A.B., Skytte, J.L., Andersson-Engels, S., Khoptyar, D.: Spectral characterisation of dairy products using photon time-of-flight spectroscopy. Journal of Near Infrared Spectroscopy **21**(5), 375–383 (2013)
15. Prince, S.J.: Computer vision: models, learning, and inference. Cambridge University Press (2012)
16. Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P.: Numerical recipes 3rd edition: The art of scientific computing. Cambridge University Press (2007)
17. Hastie, T., Tibshirani, R., Friedman, J., Hastie, T., Friedman, J., Tibshirani, R.: The elements of statistical learning, 2nd edn. Springer (2009)
18. Mie, G.: Contribution to the optical properties of turbid media, in particular of colloidal suspensions of metals. Ann. Phys. (Leipzig) **25**, 377–452 (1908)
19. Claesson, O., Nitschmann, H.: Optical investigation of the rennet clotting of milk. Acta Agriculturae Scandinavica **7**(4), 341–360 (1957)
20. O'Callaghan, D.J., O'Donnell, C., Payne, F.: A comparison of on-line techniques for determination of curd setting time using cheesemilks under different rates of coagulation. Journal of Food Engineering **41**(1), 43–54 (1999)