Machine Vision and Applications (2014) 25:327–343
DOI 10.1007/s00138-013-0512-2

ORIGINAL PAPER

Hyperspectral imaging based on diffused laser light for prediction of astaxanthin coating concentration

Martin Georg Ljungqvist · Otto Højager Attermann Nielsen · Stina Frosch · Michael Engelbrecht Nielsen · Line Harder Clemmensen · Bjarne Kjær Ersbøll

Received: 1 September 2012 / Revised: 21 February 2013 / Accepted: 10 April 2013 / Published online: 9 May 2013

Abstract We present a study on predicting the concentration level of synthetic astaxanthin in fish feed pellet coating using multi- and hyperspectral image analysis. This was done in parallel using two different vision systems. A new instrument for hyperspectral imaging, the SuperK setup, using a super-continuum laser as the light source was introduced. Furthermore, a parallel study with the commercially available multispectral VideometerLab imaging system was performed. The SuperK setup used 113 spectral bands (455–1,015 nm), and the VideometerLab used 20 spectral bands (385–1,050 nm). To predict the astaxanthin concentration from the spectral image data, the synthetic astaxanthin content in the pellets was measured with the established standard technique; high-pressure liquid chromatography (HPLC). Regression analysis was done using partial least squares regression (PLSR) and the sparse regression method elastic net (EN). The ratio of standard error of prediction (RPD) is the ratio between the standard deviation of the reference values and the prediction error, and for both PLSR and EN both devices gave RPD values between 4 and 24, and with mean prediction error of 1.4–8.0 parts per million of astaxanthin concentration. The results show that it is possible to predict the synthetic astaxanthin concentration in the coating well enough for quality control using both multi- and hyperspectral image analysis, while the SuperK setup performs with higher accuracy than the VideometerLab device for this particular problem. The spectral resolution made it possible to identify the most significant spectral regions for detection of astaxanthin. The results also imply that the presented methods can be used in general for quality inspection of various coating substances using similar coating methods.

Keywords Hyperspectral imaging · Multispectral imaging · Spectral imaging · Image analysis · Diffused laser · Partial least squares · Elastic net · Astaxanthin · Pellet coating

M. G. Ljungqvist (✉) · O. H. A. Nielsen · L. H. Clemmensen · B. K. Ersbøll
Department of Informatics and Mathematical Modelling, Technical University of Denmark, Kgs. Lyngby, Denmark
e-mail: malj@imm.dtu.dk

S. Frosch · M. E. Nielsen
Division of Industrial Food Research, National Food Institute, Technical University of Denmark, Kgs. Lyngby, Denmark

1 Introduction

Astaxanthin is a naturally occurring carotenoid with a high antioxidant activity essential for growth and survival and it is important for the development of colour in salmonid fishes [35]. The primary use of astaxanthin within aquaculture is as a feed additive to ensure that farmed salmon and trout have similar appearance to their wild counterparts [38]. The colour appearance of fish products is important for customers as a quality indicator [16,29,34]. Astaxanthin is very expensive [2] and therefore optimisation of its use in fish feed production is of importance.

Synthetic astaxanthin is more easily available and costs slightly less than natural astaxanthin and is therefore used more often in industry. However, there is a demand for natural astaxanthin for the organic salmonid fish market, where natural astaxanthin is mandatory.

Today, chemical measurement of astaxanthin is done by high-pressure liquid chromatography (HPLC). HPLC is a well-established technique for measuring synthetic astaxanthin content in fish oil. However, when measuring astaxanthin concentration in fish feed pellets, an additional step for extraction of the oil is necessary, and this is why

measurements of astaxanthin from fish feed pellets are less accurate and more labour intensive.

An automatic vision system for at-line pigment quality control of astaxanthin coating concentration level would be of great benefit to the industry, in relation to both process control and process optimisation.

The aim of this study is to investigate the possibility of predicting the concentration level of synthetic astaxanthin coating on aquaculture feed pellets by spectral imaging. Since HPLC is used for measuring astaxanthin coating content in the industry, we used this as a reference method.

Spectral imaging is called multispectral when using a small number of spectral bands (e.g. less than about 50), and hyperspectral when using a large number of spectral bands (e.g. hundreds).

Previous studies on feed pellet monitoring by spectral analysis include near infra-red (NIR) reflectance spectroscopy for classification of feed material and feed pellets by Fernández-Ahumada et al. [12], and predicting chemical information in pharmaceutical pellet core and coating using NIR imaging by Sabin et al. [31].

Previous work on multi- and hyperspectral image analysis of astaxanthin coating by Ljungqvist et al. [22,23] has shown promising results for screening of the concentration level. However, those studies did not use hyperspectral imaging in the visual part of the spectrum. In [23] only the NIR range was analysed, and in [22] only 20 wavelengths were analysed in the visual and parts of the NIR range. Neither of them used chemical measurements for validation of the astaxanthin coating level.

For the previous work using the visual part of the spectrum [22], the spectral bands were located at predefined wavelengths due to instrument setup, chosen without the knowledge of the particular problem. The few spectral bands used may not be the ones that give the greatest ability to quantify the contents of astaxanthin. Thus, a more detailed study is called for, and here we report the characteristics for a new imaging method based on diffused laser light with more spectral bands. The study is focused on the visual region of the electromagnetic spectrum due to function of astaxanthin as a pigment.

Vision systems have previously been implemented for quantisation of chemical contents, and a number of light-source options exist. In this paper, we present the use of a super-continuum light source for full-field illumination. The super-continuum laser, combined with an acousto-optical tunable filter (AOTF) provides a broadband tunable light source. This form of light source is often used for confocal microscopy [32], fluorescence lifetime imaging [24], and measurement of subsurface laser scattering (SLS) [26,27]; also known as diffuse reflectance. To apply the light-source for full-field illumination, the small beam from the AOTF box is simply expanded and made diffuse.

A parallel study with a commercially available multispectral system called VideometerLab was also performed. The performance of this device has been validated for similar surface chemistry applications [6–9,11,15,17,21].

Since the hyperspectral imaging device records more spectral bands than a multispectral device, it would give more detailed information for measuring the astaxanthin coating concentration.

The presented work thus investigates both the possibility of examining astaxanthin contents by hyperspectral image analysis, and a comparison of the two modalities (multi- and hyperspectral imaging) for astaxanthin prediction.

2 Materials and methods

2.1 Material

The pellets used in this study were produced for the purpose of this experiment, and the recipe was based on normal commercial fish feed for salmonid fish. The pellets had the approximate production diameter of 4.5 mm, and were coated with fish oil.

An extruder machine was used for the pellet production. The feed material was extruded through a die plate with holes of a specific diameter which determined the diameter of the pellets. On the other side of the disk, there was a set of rotating knives that cut the material into shorter, cylinder-shaped pellets.

The synthetic astaxanthin used was cold water dispersible (BASF SE, Germany), and in total seven different levels of synthetic astaxanthin concentration were added to the fish oil coating. The highest synthetic astaxanthin level in fish oil was 100 parts per million (ppm), and then the oil was diluted so that the concentration level became half of the original. This was repeated to achieve the seven nominal levels of synthetic astaxanthin concentration in the fish oil coating, where the last level was 0 ppm, see Table 1.

Fish oil in itself typically contains a small amount of natural astaxanthin, however, this is assumed to be less than 1 ppm and is here referred to as a coating of 0 ppm concentration. Astaxanthin is commonly measured in ppm, and it is measured in mass so here ppm corresponds to mg/kg.

Between production and image acquisition, the pellets were stored at 2 °C in a dark environment for 2 months. They were stored in plastic bags where the oxygen had been pumped out and the bags were filled with nitrogen to minimise the oxidation process and quality reduction during storage.

The spectral reflection of the pellets is a mix of the pellet compound (recipe) and the reflection of astaxanthin. The light captured by each pixel is thus assumed to be a linear combination of two main components; the pellet compound

Table 1 Data overview of the seven pellet groups (nominal concentration levels) and the number of images taken (of different samples) with VideometerLab and SuperK, respectively

	Synthetic astaxanthin concentration (ppm)						
Levels	0	3.125	6.25	12.5	25	50	100
	Number of images						
VideometerLab	30	30	30	30	30	30	30
SuperK	10	10	10	10	10	10	10

Added synthetic astaxanthin to the fish oil coating in ppm

and the astaxanthin coating. Due to the production method, the coating is neither evenly distributed among the pellets, nor uniformly on each pellet.

In the extrusion process, parts of the astaxanthin coating will go inside the pellet, while we only measure reflection from the surface. However, the industry is interested in the total amount of astaxanthin in the pellets. Therefore, the surface reflection is assumed to be linearly related to the total amount of astaxanthin in the pellets.

It is assumed that most of the quantity of synthetic astaxanthin on each pellet can be estimated by measuring the coating surface of each pellet. For practical use, however, it is not interesting to estimate the quantity on each single pellet, but rather on a larger amount of pellets, and calculate an ensemble average.

2.2 Equipment

2.2.1 SuperK

The hyperspectral imaging system consists of four parts: light source, spectral filter, diffuse filter and camera. The illumination system is based on a SuperK Extreme (NKT Photonics A/S, Denmark) super-continuum white light laser producing a quasi-continuous output. The super-continuum light is filtered using a SuperK Select (NKT Photonics A/S, Denmark), where an AOTF is used for spectral filtering of the beam. The combined light source and filtering box provides a wavelength tunable laser beam delivered in a photonic crystal fibre (PCF) by a Fiber Delivery System (NKT Photonics A/S, Denmark). The combined system provides 0.5–6.5 mW.

In combination, this system provides light in the visual and NIR region, ranging from 455 to 1,015 nm. A step size of 5 nm was used as the spectral resolution, resulting in 113 spectral bands. The bandwidth grows linearly as a function of wavelength; at 500 nm it is 3.5 nm, at 900 nm it is 14 nm.

The spectrally filtered light from the Fiber Delivery System forms a Gaussian fundamental transverse electromagnetic (TEM_{00}) beam. This is transformed to illuminate a square area below the camera using an engineered diffuser (ED1-C20-MD, Thorlabs, Sweden). This diffusing method has a high power transmission onto the sample.

The illumination does not form a perfectly uniform intensity distribution; it produces a gradient due to the projection caused by the oblique incidence of the beam onto the camera field of view. It also produces a short distance intensity fluctuation. To minimise the latter effect, the Gaussian beam delivered after the AOTF box is expanded using a -50 mm focal length negative lens (LC1439, Thorlabs, Sweden). By illuminating the engineered diffuser, the short distance intensity fluctuations are minimised and become insignificant for this application.

Image capturing is done using a Grasshopper GRAS-20S4M grey-scale charge-coupled device (CCD) camera (Point Grey Research Inc, Canada), which uses a 12-bit analogue to digital converter (ADC) with a 16-bit output. The image resolution is 1,600 $\times$ 1,200 pixels with a pixel size of approximately 0.028 $\times$ 0.028 millimetres. For an overview of the SuperK setup see Fig. 1.

2.2.2 VideometerLab

To compare the result of the SuperK setup, the commercially available multispectral VideometerLab (Videometer A/S, Hørsholm, Denmark) was also used for image acquisition. It uses 20 wavelengths distributed over the ultra-violet A (UVA), visual and NIR region: 385, 430, 450, 470, 505, 565, 590, 630, 645, 660, 700, 850, 870, 890, 910, 920, 940, 950, 970, 1,050 nm.

This system uses a Point Grey Scorpion SCOR-20SOM grey-scale CCD camera. The objects of interest are placed inside an integrating sphere (Ulbricht sphere) with uniform diffuse lighting from light sources placed around the rim of the sphere [5]. All light sources are light-emitting diodes (LED) except for 1,050 nm, which is a diffused laser diode.

The curvature of the sphere and its matt-white coating ensure a uniform diffuse light so that specular effects are avoided and the amount of shadow is minimised. The device is calibrated radiometrically with a following light and exposure calibration according to the National Institute of Standards and Technology (NIST). The system is also

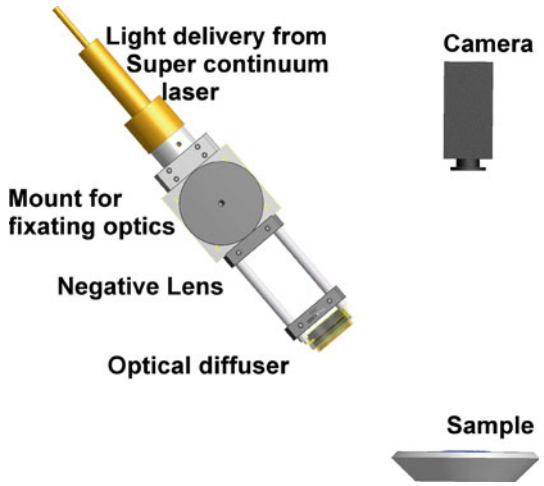

Fig. 1 Overview of the SuperK setup showing the supercontinuum light delivery, negative lens, optical diffuser, camera, and sample position

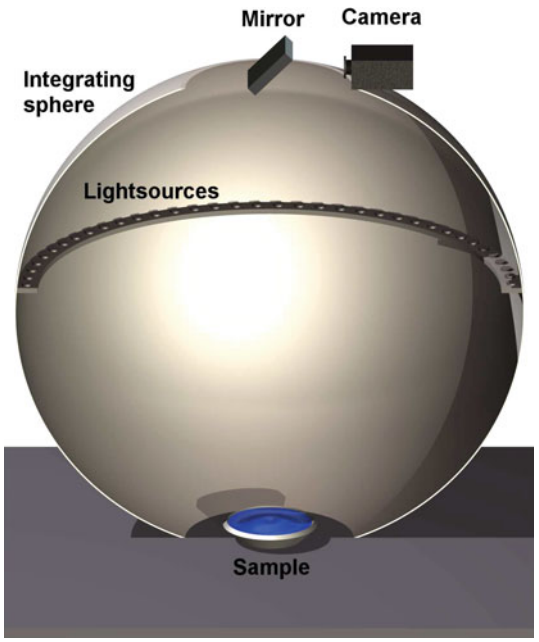

Fig. 2 The VideometerLab setup showing a cross section of the integrating sphere, light sources, camera, and sample position

geometrically calibrated to ensure pixel correspondence for all spectral bands [13]. For an overview of the VideometerLab setup see Fig. 2.

VideometerLab has the advantage that the intensity is calibrated with respect to the sensitivity of the CCD and the intensity of the light sources, which means that the resulting reflection spectrum can be compared with, e.g. a spectroscopy spectrum.

The image resolution is $1{,}200 \times 1{,}200$ pixels. In this situation, one pixel represents approximately 0.072×0.072 millimetres. The Scorpion camera has a 12-bit ADC, and the system uses 8-bit data output from the camera. After calibration correction, the reflectance intensity output is at 32-bit precision.

2.2.3 Spectroscopy

To explore the spectral properties of astaxanthin further, and to assist in the interpretation of the spectral image results, a spectrometer was used in the visual and NIR range. Absorption spectra of synthetic astaxanthin in a solution of fish oil along with plain fish oil were recorded using a NIRSystems 6500 absorption spectrometer (Foss NIRSystems Inc, USA) with a spectral resolution of 2 nm. The absorption spectra were transformed to reflection values using the standard relation $A = -\log(R)$, where A is the absorption value and R is the reflection value.

2.2.4 High-pressure liquid chromatography

To calibrate the hyperspectral imaging prediction method, an HPLC analysis of the synthetic astaxanthin concentration in pellets was performed. By analysing samples from each concentration level, we could estimate the average astaxanthin content in the pellets, which can be used to validate the nominal levels used in the production. To reduce the effect of the analysis method, we used two independent HPLC measurements from different parties.

The HPLC analysis was done at the National Food Institute, Division of Industrial Food Research at the Technical University of Denmark (Lyngby, Denmark) using an Agilent 1100 series HPLC (Agilent Technologies, Palo Alto, CA, USA), equipped with a UV diode array detector. The oil was extracted from the pellets using acetone and homogenised to a concentrate, which was analysed for synthetic astaxanthin content, according to the modified protocol of Bligh and Dyer [3].

A fraction of the lipid extract was evaporated under nitrogen and redissolved in 2 mL of *n*-heptane before injection. Astaxanthin content was determined after injection of an aliquot (50 µL) of the *n*-heptane fraction into a LiChrosorb Si60-5 column (100 mm × 3 mm, 5 µm) equipped with a Cromsep Silica (S2) guard column (10 mm × 2 mm; Chrompack, Middelburg, The Netherlands) and eluted with a flow of 1.2 mL/min using *n*-heptane/acetone (86:14, v/v) and detection at 470 nm. Concentrations of astaxanthin were calculated using authentic standards from Dr. Ehrenstorfer GmbH (Augsburg, Germany).

Furthermore, HPLC analysis was also made at Eurofins A/S (Galten, Denmark), which is a commercial laboratory.

2.3 Image acquisition

Images of petri dishes (plastic, diameter of 9 cm) filled with pellets were captured using both the SuperK setup and the VideometerLab. For each concentration level, 10 images of different pellets were captured using both SuperK and VideometerLab, then an additional 20 images were captured with VideometerLab, see Table 1. In total, 70 SuperK pellet images and 210 VideometerLab pellet images were captured over two consecutive days.

The concentration level sequence was randomised, and samples were interleaved two at a time. The pellets were at normal room temperature during image acquisition.

A filled petri dish resembles the rapid inspection that the industry would desire for this application.

The pellet cluster inside the petri dish in each image was segmented from the light-grey background using a grey-scale threshold, using the band of 500 nm for the SuperK images and the band of 470 nm for the VideometerLab images. The threshold segmentation was complemented with the morphological methods erosion and dilation using a disk as structuring element [33]. Furthermore, the topmost layer of pellets was segmented to remove parts with less light, and in this way also avoid some of the granulometry information in the image. Since the SuperK images contained some specular reflections due to the direct lighting, the strongest specular effects were also removed by a threshold for both the VideometerLab and SuperK images.

The mean spectrum of the pixels in each segmented image was used as samples. In this way, the impact of the pattern from the diffuse filter used by the SuperK setup was reduced and was assumed not to impact the results of the image analysis.

Standard red–green–blue (sRGB) colour image representations of the VideometerLab images for this paper were done by multispectral colour-mapping using penalised least square regression described in Dissing et al. [10]. Since the SuperK setup is not calibrated towards the CCD, sensitivity-standardised colour mapping was not possible, instead non-standard RGB images were made to visualise the images.

2.4 Data analysis

The number of samples is denoted by n and the number of variables (the wavelengths) is denoted by p. The stored data of n samples and p variables are denoted as matrix $\mathbf{X}$. The ground truth reference values (concentration level) are stored in vector $\mathbf{y}$ with length n. The predicted (estimated) value of $\mathbf{y}$ is denoted $\hat{\mathbf{y}}$.

For ground truth reference values, the nominal values in Table 1 were used.

For the VideometerLab data p is 20 and n is 210. To compare the VideometerLab data with the SuperK data, the 70 images corresponding to the same petri dishes of pellets for both methods were used in the analyses; $n = 70$. For the SuperK data p is 113 and n is 70, this results in an ill-posed problem with more variables than observations ($p \gg n$).

Two different regression methods were used to estimate the concentration level of synthetic astaxanthin in the pellet coating. Both methods produce linear prediction models, but they calculate the models in different ways, as described below.

All image analyses and statistics were carried out using Matlab 7.9 (The Mathworks Inc., Natick, MA, USA).

2.4.1 Principal component analysis

The multivariate data from the images were analysed using principal component analysis (PCA) for exploratory purposes [14]. PCA is an unsupervised method, and the most optimal method with respect to maximising the variance in the data [18]. If the relation of interest contains large variations, then PCA is a good method for analysing the data. The pre-processing method, called standard normal variate (SNV) [30], was used, followed by PCA. SNV is performed by subtracting the mean from each sample, and normalising using the standard deviation (SD) of the sample spectrum.

2.4.2 Model selection and validation

To calibrate the statistical model parameters, a calibration set of 70 % of the samples was used (n_c). Then, the chosen model was validated on the remaining 30 % of the samples (n_v). For the VideometerLab data with a total of 210 samples, $n_c = 147$ and $n_v = 63$, and for the VideometerLab data using 70 samples, $n_c = 49$ and $n_v = 21$. For the SuperK data with a total of 70 samples, $n_c = 49$ and $n_v = 21$.

The calibration and validation set was chosen randomly, but with all concentration levels present in both. The same corresponding samples for both devices were present in both the calibration and the validation set.

For parameter calibration, two different methods were used in parallel. First, the leave-one-out cross-validation (LOOCV) method was used on the calibration set, where each sample is used as validation once.

Second, what can be called a group-fold cross-validation (GFCV) was used both for parameter calibration and to investigate how the prediction generalises on unseen concentration levels (validation). The GFCV was performed so that all samples with one concentration level, a group, were left out during training of the model, and then the left out group was used for prediction.

Two nested GFCV:s were used, one for calibrating a PLSR model, and one for the final validation and generalisation. For every iteration, a PLSR model was calibrated using GFCV on six levels, and then the validation was done on the left

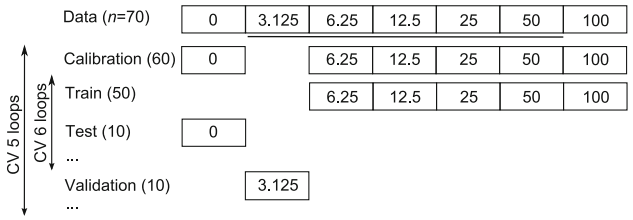

Fig. 3 Schematic overview of the group-oriented cross-validation GFCV for calibration and validation of the prediction. The GFCV is used both for calibration and for generalisation validation. The data contain seven different concentration levels, groups, and the middle five of these are used for generalisation validation. Each group consists of ten samples. Calibration is performed on six groups at a time. One group at a time is left out for validation. Synthetic astaxanthin concentration in ppm

out level. For each iteration, one of the five middle levels (3.125, 6.25, 12.5, 25, 50 ppm) was left out at a time to see how a model predicts on an unseen concentration level. The highest and lowest concentration levels were not used for the validation/generalisation part since we are not interested in extrapolation of the model. This procedure renders a result in terms of a pooled prediction error of the five group tests. For an overview of the GFCV scheme see Fig. 3.

2.4.3 Partial least squares regression

The first method for analysing the concentration level was the partial least squares regression (PLSR) method [18,37]. The data were mean centred, and the number of components to be used in the PLSR model was decided using LOOCV, and likewise GFCV, on the calibration set while minimising the mean of the residual sum of squares (RSS), see Eq. 1. A modified version of the 'one standard error rule' [18] for selection was used: The least number of components with RSS value inside the range of two SDs of the lowest value of RSS was selected.

For the VideometerLab data, the number of components (factors) n_f tested in the calibration step was varied from 1 to 20. For the SuperK data containing 113 variables and a calibration set of 49 samples; 47 was used as the maximum number of components tested.

The chosen value of n_f was used on the calibration set and the resulting prediction model was then validated on the validation set.

The coefficient of determination R^2 is a measure of how much variation is explained by the model and was calculated for the prediction of the validation set. R^2 is basically the ratio of the RSS and the total sum of squares (TSS), see Eq. 1.

$$R^2 = 1 - \frac{\text{RSS}}{\text{TSS}} = 1 - \frac{\sum_{n_v}(\mathbf{y} - \hat{\mathbf{y}})^2}{\sum_{n_v}\left(\mathbf{y} - \frac{1}{n_v}\sum_{n_v}\mathbf{y}\right)^2}. \quad (1)$$

Furthermore, the ratio of the standard error of prediction and standard deviation (RPD) was calculated as a measure of how well the model predicts. The RPD is the ratio between the SD of the original data $\mathbf{y}$ (the reference values) and the standard error of prediction (SEP), see Eqs. 2 and 3.

$$\text{SEP} = \text{RMSEP} = \sqrt{\frac{1}{n_v}\sum_{n_v}(\mathbf{y} - \hat{\mathbf{y}})^2}. \quad (2)$$

$$\text{RPD} = \frac{\text{SD}}{\text{SEP}}. \quad (3)$$

The SEP is equal to the root mean square error of prediction (RMSEP). An RPD value of 1.0 means that the model cannot predict accurately, since this means that the mean error is equal to the SD of the reference values. An RPD value higher than 2.5 is considered satisfactory for screening, and the values of 5–10 are adequate for quality control [39].

To further see what range of the spectrum is relevant for the prediction of astaxanthin concentration, we used interval PLS (iPLS) regression where the spectrum is divided into a number of regions and PLSR is performed on each region individually [28]. While the VideometerLab data has too few spectral bands for this method, we can investigate what parts of the SuperK spectrum give best prediction. We used 20 intervals and a PLSR model was calibrated for each of them.

2.4.4 Elastic net regression

To identify which wavelengths are of most interest, the elastic net (EN) was also used for regression analysis [41]. The EN is a sparse statistical method and performs variable selection while doing regression. In contrast to PLSR it does not use linear combinations of all variables, it uses only a few variables which are found important; non-zero coefficients. In this way, EN excludes variables that do not contribute to the result and thus are potential noise.

EN tends to select variables that are correlated with each other, and this is suitable for spectral data with intrinsic correlation. This grouping effect would make the result suitable for an optical filter implementation.

EN needs two model parameters: λ_1 for the L_1 norm for determining the number of non-zero coefficients, and λ_2 for the Euclidean L_2 norm for regularisation. The regularisation is suitable for the ill-posed problem. The regression model consists of the variable weights (coefficients) in $\boldsymbol{\beta}^{\text{en}}$, see Eq. 4.

$$\beta_j^{\text{en}} = \underset{\beta_j}{\text{argmin}}(\|\mathbf{y} - \mathbf{X}\beta_j\|_2^2 + \lambda_2\|\beta_j\|_2^2 + \lambda_1\|\beta_j\|_1). \quad (4)$$

The estimated coefficients are then multiplied by $(1+\lambda_2)$ to get the final EN solution.

The two parameters were selected using LOOCV on the calibration set. The λ_1 parameter steers the selection of variables and was calculated so that the number of selected variables was varied from 1 to 20. The λ_2 parameter was varied with 12 logarithmic steps from 10^{-7} to 10. The data matrix **X** was normalised, and the reference values **y** were centred for each calculation of the EN.

RSS was used to find the optimal parameter set. If more than one combination of the number of selected variables and λ_2 was found to give the best calibration result, then the lowest number of selected variables and the highest value of λ_2 were used, giving the least complex model.

The implementation by Sjöstrand et al. [36] was used for calculations of the EN.

3 Results

3.1 SuperK power stabilisation

To estimate the acquisition reproducibility, the same sample was repeatedly imaged for ∼11 h, and the mean image intensity was measured. To compare across wavelengths, we report the variation relative to the mean intensity of the given wavelength; the ratio denoted as a percentage (%). The resulting SDs are presented in Fig. 4 showing the image acquisition reproducibility.

Compared to the precision that is desired in estimating astaxanthin coating concentration (∼ 1 %), we suspect that the wavelengths below 470 nm will be too noisy for modelling the concentration level. For the regression models this will automatically mean a reduced weighting for these bands.

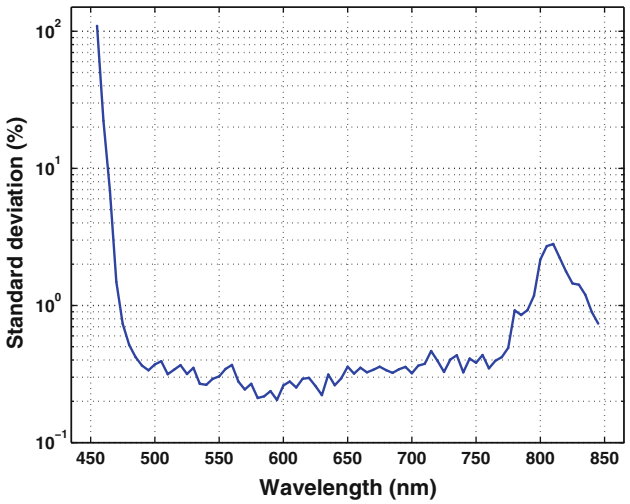

Fig. 4 SuperK stability measurement. Standard deviation of the mean image intensities (%) of the power-normalised image data for each wavelength

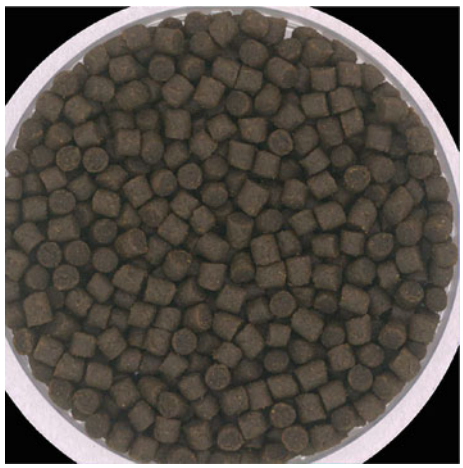

Fig. 5 SRGB representation of a VideometerLab image of pellets coated with synthetic astaxanthin of 50 ppm

Depending on which wavelength is chosen during modelling, and since we only apply linear models, the reported SD as a function of wavelength indicates the precision of ∼0.3 % as a limit of the acquisition process itself.

3.2 Chemical measurement

The results from the two independent HPLC measurements of the synthetic astaxanthin coating concentration both show lower values than the nominal values, and DTU shows lower values than Eurofins. However, both HPLC measurements are linear to the nominal values, see Table 2, which means that the nominal values are valid as reference values for the regression.

3.3 Image analysis

Background and dark sample parts were segmented out using threshold and morphological operations. Example results of a segmented VideometerLab image can be seen in Figs. 5 and 6, as well as for SuperK in Figs. 7 and 8. RGB versions of the captured images for all seven concentration levels visualising the differences in pigment concentration can be seen in Fig. 9 for VideometerLab, and in Fig. 10 for the SuperK images.

The spectrometer spectra of fish oil with added synthetic astaxanthin and plain fish oil are presented in Fig. 11, which shows that the largest separation is in the range of 490–610 nm.

Investigating the reflection spectra of the pellets coated with synthetic astaxanthin shows similar characteristics to the spectroscopy results of astaxanthin in fish oil, see Fig. 12. Looking at the difference of the concentration levels' reflection spectra in Figs. 13 and 14 reveals that the largest difference between concentration levels is around 505 nm for the

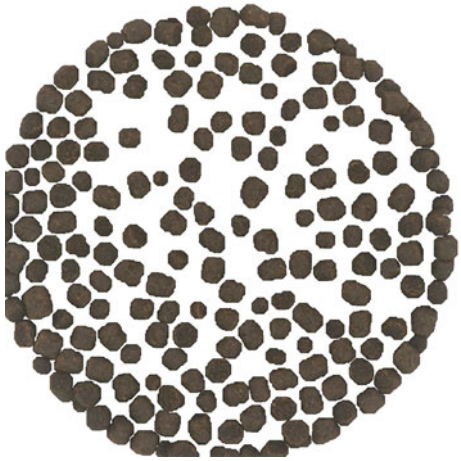

Fig. 6 SRGB representation of the segmented version of the VideometerLab image in Fig. 5 with pellets coated with synthetic astaxanthin of 50 ppm

Fig. 7 RGB representation of a SuperK image of pellets coated with synthetic astaxanthin of 50 ppm

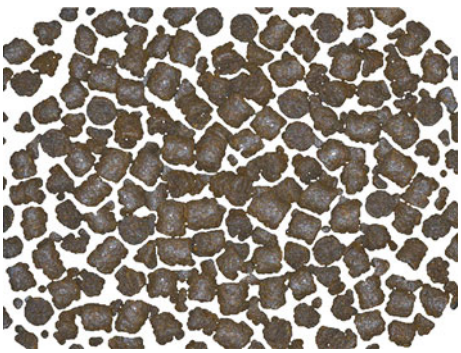

Fig. 8 RGB representation of the segmented version of the SuperK image in Fig. 7 with pellets coated with synthetic astaxanthin of 50 ppm

VideometerLab data, and between 500 and 520 nm for the SuperK data. This corresponds well with regions of difference in the reflection spectrum in Fig. 11.

From the explorative analysis of the concentration level of synthetic astaxanthin coating using PCA, score scatter plots using all combinations of PC1 to PC6 were investigated (results not shown). PC2 seems to describe the concentration scale well for the SuperK data, and PC2 in combination with PC3 for the VideometerLab data. However, both show some small overlap—mostly for the lower concentration levels, see Figs. 15 and 16. This concludes that the variation between the concentration levels is the second largest variance in the image data (PC2), next to the total intensity variation (PC1), which corresponds well with the two spectra mentioned above.

Calibrating the PLSR model by LOOCV gives an optimum of three components for the VideometerLab data, and seven components for the SuperK data, see Fig. 17. It also shows that the calibration error drops significantly around three components for both modalities.

Regression analysis of the synthetic astaxanthin coating concentration level using the whole images as samples shows good results. PLSR using LOOCV shows an R^2 value of prediction above 0.94 for VideometerLab and 0.99 for SuperK, see Table 3, where VideometerLab uses three PLSR components and SuperK uses seven PLSR components.The EN using LOOCV has an R^2 around 0.97 for VideometerLab and 0.99 for SuperK, see Table 3, whereas the VideometerLab data uses eight non-zero coefficients, and the SuperK data uses 19 coefficients.

For the VideometerLab data, the RPD is between 4 and 9 for both PLSR and EN, while for the SuperK data the RPD value is almost 24 for PLSR and 21 for the EN. The mean prediction error is 7.6 ppm for VideometerLab using 210 samples, and 8.0 ppm when using 70 samples. For the SuperK data using 70 samples, the mean prediction error is 1.4 ppm. The synthetic astaxanthin coating concentration prediction using PLSR and EN can be seen in Figs. 18 and 19, illustrating that the predicted levels correlate well with the nominal levels. For the SuperK data, three PLSR components were also tested resulting in an R^2 of 0.97, SEP of 3.8, and an RPD of 9, see Table 3.

The RPD for different numbers of PLSR components can be seen in Fig. 20, where it shows that RPD above 5 is achieved for SuperK using only two components, and using four components for VideometerLab.

The PLSR components' weights show the contribution of different spectral regions to the model, and thus their significance. Similar to the difference from the grand mean, Figs. 13 and 14, we rediscover the importance of the blue and green regime between 450 and 590 nm, but see distinctions also in the NIR measurements for the two instruments, see Figs. 21 and 22. For the VideometerLab data the largest PLSR weights are at 385 and 505 nm, while for the SuperK data the largest PLSR weights are around 550 nm and around 860 nm. The importance of the visual range of the spectrum was confirmed by the iPLS regression of the SuperK data returning the lowest LOOCV calibration error around 545–575 nm, see Fig. 23.

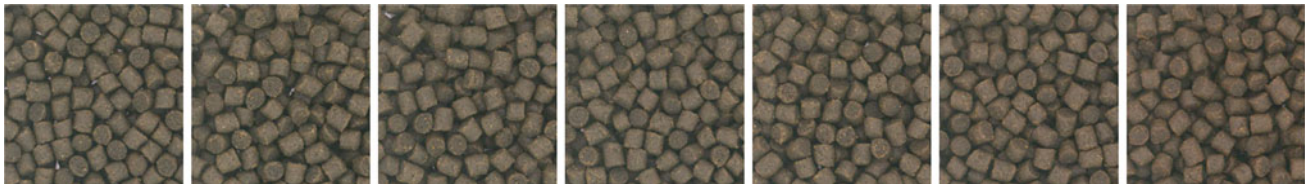

Fig. 9 SRGB representation of cropped VideometerLab images of the pellets with seven different levels of synthetic astaxanthin coating concentration. From left to right: 0, 3.125, 6.25, 12.5, 25, 50, 100 ppm

Fig. 10 RGB representation of cropped SuperK images of the pellets with seven different levels of synthetic astaxanthin coating concentration. From left to right: 0, 3.125, 6.25, 12.5, 25, 50, 100 ppm

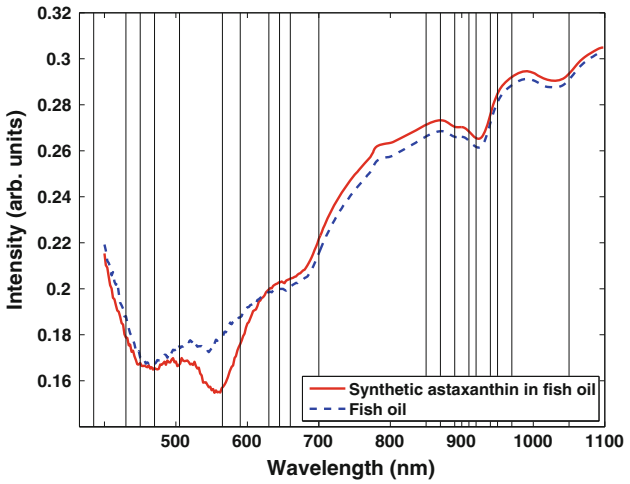

Fig. 11 Mean reflection spectra of synthetic astaxanthin coating in fish oil, as well as plain fish oil, using a spectrometer. No pellets were used here. The bands of the VideometerLab are marked by vertical lines

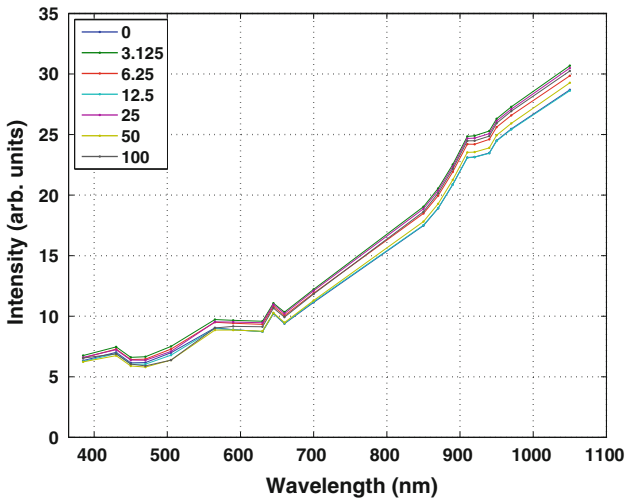

Fig. 12 Mean reflection spectra of the different concentration levels of synthetic astaxanthin in pellet coating, using image mean of the VideometerLab images, $n = 70$

The EN regression coefficients $\boldsymbol{\beta}_{\text{en}}$ that constitute the prediction model show similar results for both VideometerLab and SuperK data, see Fig. 24. For the VideometerLab data the largest coefficient is at 565 nm, and the second largest at 505 nm. For the SuperK data the largest coefficient is at 550 nm. Both systems show similar clusters of spectral bands with high emphasis in the visual regime.

The results also show that using the 16 bands that are the same for both the VideometerLab and the SuperK setup (470, 505, 565, 590, 630, 645, 660, 700, 850, 870, 890, 910, 920, 940, 950, 970 nm) gives acceptable results for quality control with an RPD value of 2.6 for VideometerLab, and 8.3 for SuperK. The VideometerLab images got an R^2 of 0.84 using 2 PLSR components, and the SuperK images got R^2 of 0.94 using 4 PLSR components, see Table 3.

By leaving out one concentration level for validation using GFCV and PLSR, it was tested how the prediction performs on unseen concentration levels. Using LOOCV, all concentration levels have been used for calibration by leaving out one sample at a time, and for comparison with GFCV here 0 and 100 ppm have been left out when performing the validation. When using GFCV all concentration levels have been used for calibration while leaving out one concentration level at a time, and all concentration levels but 0 and 100 ppm have been used for validation, one at a time. We see that the pooled SEP of GFCV is larger than the corresponding SEP of LOOCV for both systems, as can be seen in Table 4. For the VideometerLab using LOOCV, the SEP is 5.5, and for GFCV the pooled SEP is 14. For the SuperK data using LOOCV, the SEP is 1.3, and for GFCV the pooled SEP is 9.6.

Table 2 HPLC measurement of the synthetic astaxanthin present in the pellets, carried out by two different parties

Nominal Values	HPLC DTU	HPLC Eurofins
0	0	0
3.125	1.90	3.07
6.25	4.36	5.88
12.5	9.56	11.6
25	19.44	24.8
50	42.64	49.3
100	90.46	95.1

Added synthetic astaxanthin to the fish oil coating in ppm

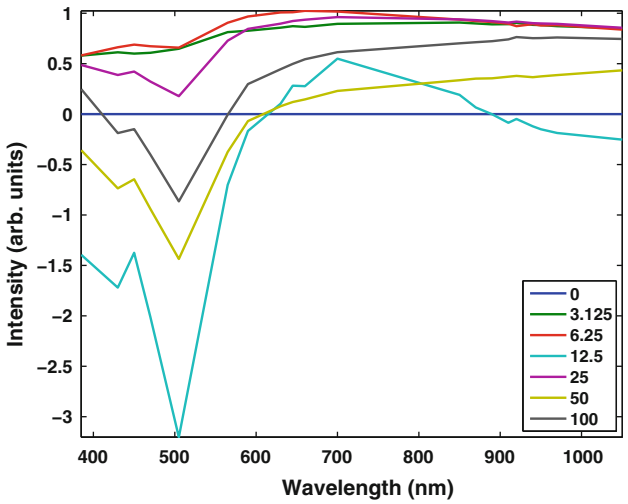

Fig. 13 The difference in reflection spectra from the concentration level of 0 ppm of the different concentration levels of synthetic astaxanthin coating, using the VideometerLab images, $n = 70$. Normalised by the SD of each level

Visualisation of the spatial distribution of synthetic astaxanthin coating on the pellets was done using the PLSR prediction result for VideometerLab and SuperK, and can be seen in Figs. 25 and 26. Spectral images from the validation set and all different concentration levels have been projected using the PLSR components. This should only be seen as a visualisation, since the PLSR models are calibrated on image mean values and then used for prediction on pixel values. The visualisations clearly show larger values for the higher concentration levels of synthetic astaxanthin.

4 Discussion

The results from the multispectral VideometerLab and the hyperspectral SuperK indicate that it is possible to predict synthetic astaxanthin coating on pellets.

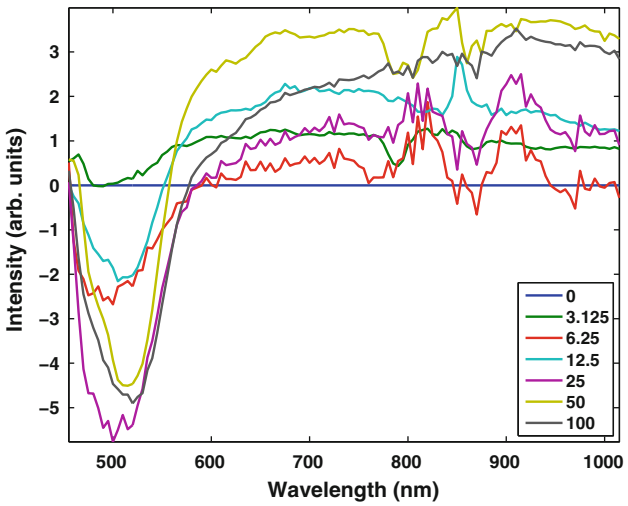

Fig. 14 The difference in reflection spectra from the concentration level of 0 ppm of the different concentration levels of synthetic astaxanthin coating, using the SuperK setup, $n = 70$. Normalised by the SD of each level

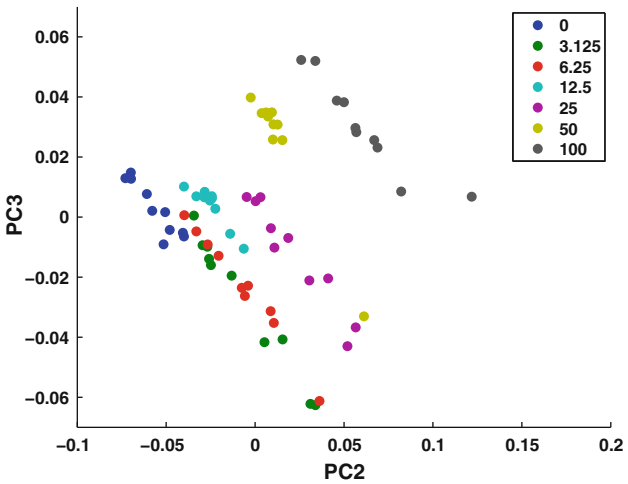

Fig. 15 Scatter plot of synthetic astaxanthin coating using PC2 and PC3 from a PCA on whole image samples using the VideometerLab, $n = 70$

The design has minimised the structure of the semi diffuse illumination used in the SuperK setup so it visually becomes insignificant. It is evident that the presented technique using an engineered diffuser provides a sufficient trade-off between a smooth uniform illumination and high power transmission.

In addition, the SuperK system's image acquisition reproducibility was investigated. The results presented in Figure 4 show low variation above 470 nm.

In industry, it is desired to know the astaxanthin coating level with an accuracy of roughly 1 ppm, which corresponds to a sensitivity of 1 % of the pixel intensity for the present study. For this reason, it was satisfactory that the reproducibility SD of the power stabilised SuperK images

Table 3 PLS and EN regression for synthetic astaxanthin coating concentration level prediction using LOOCV. PLS regression was also done using only the 16 bands that are in common for VideometerLab and SuperK

Device	Samples n	Method	Comp. coef.	R^2 calibr.	R^2 pred.	SEC	SEP	RPD pred.
VideometerLab								
	210	PLS	3	0.9483	0.9479	7.6916	7.6180	9.1292
	70	PLS	3	0.9598	0.9421	6.9785	8.0281	4.2588
	210	EN	8	0.9729	0.9728	5.4904	5.5053	6.1094
	70	EN	8	0.9853	0.9702	4.0426	5.7571	5.9387
SuperK								
	70	PLS	7	0.9990	0.9982	1.1345	1.4278	23.9466
	70	PLS	3	0.9788	0.9872	5.0698	3.7728	9.0622
	70	EN	19	0.9989	0.9976	1.1093	1.6209	21.0926
VideometerLab-16								
	70	PLS	2	0.8441	0.8432	13.5960	13.2122	2.5877
SuperK-16								
	70	PLS	4	0.9850	0.9848	4.3091	4.1108	8.3170

Synthetic astaxanthin concentration in ppm

Table 4 Generalisation test with GFCV using PLS regression for synthetic astaxanthin coating concentration level prediction. The pooled SEP of GFCV is compared to the corresponding prediction error of LOOCV, where only the middle concentration levels (3.125, 6.25, 12.5, 25, 50 ppm) have been used for validation for both methods

Device	CV type	Samples n	Comp. n_f	SEP
VideometerLab				
	LOOCV	70	3	5.4971
	GFCV	70	2	14.0364
SuperK				
	LOOCV	70	7	1.3240
	GFCV	70	1.6	9.5813

The mean number of PLSR components used for GFCV is presented. Synthetic astaxanthin concentration in ppm

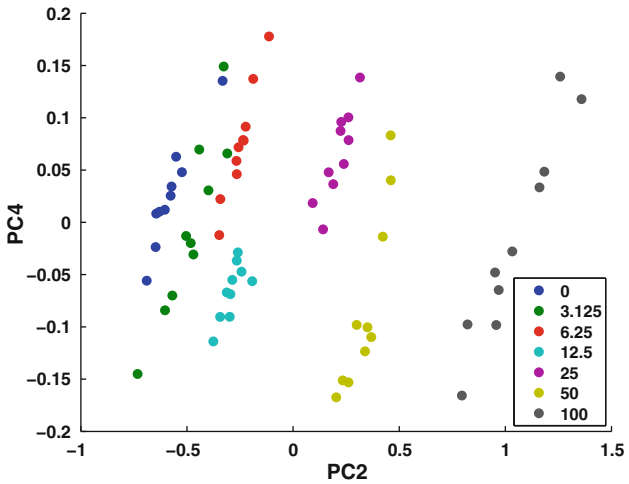

Fig. 16 Scatter plot of synthetic astaxanthin coating using PC2 and PC4 from a PCA on whole image samples using the SuperK, $n = 70$

taken over long acquisition times (11 h) proved to be below 1 %, see Fig. 4.

For each image acquisition, the two systems examined different numbers of pellets due to the difference of field of view. However, this difference was not significant since a large number of pellets were investigated in each image.

Previous studies of astaxanthin [1,4,40] found absorbance peaks of astaxanthin of around 450–505 nm and secondary peaks of around 500–600 nm for various solvents, as well as at 870 nm. This corresponds with the spectrometer results seen in Fig. 11.

To continue the data exploration, we have presented the spectral difference of the different synthetic astaxanthin coating concentration levels, seen in Figs. 13 and 14. For both modalities, the largest discrimination between data was present in the ∼500–550 nm wavelength range. This is in good agreement with the difference in spectroscopy spectra, presented in Fig. 11, of fish oil with and without synthetic astaxanthin. Since astaxanthin is a pigment, it makes sense that the visual range of the spectrum is of importance in the results.

The two vision systems presented very similar characterisations of the samples. Spectral response from both systems shows a clear separation between concentration levels for the wavelengths of about 510–530 nm, and also along the entire spectrum. SuperK showed more distinction between the small concentration levels, which can be seen both in the difference spectra and the PCA plot, see Figs. 14 and 16.

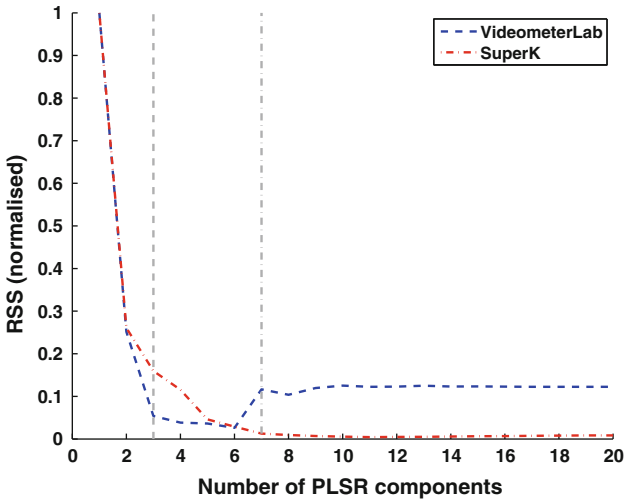

Fig. 17 PLSR calibration using LOOCV RSS, using the calibration data set of VideometerLab and SuperK respectively, $n_c = 49$, $n = 70$. Minimum of RSS with the least number of components within two SD was found at three components for VideometerLab, and at seven components for SuperK. The first 20 components are shown for SuperK

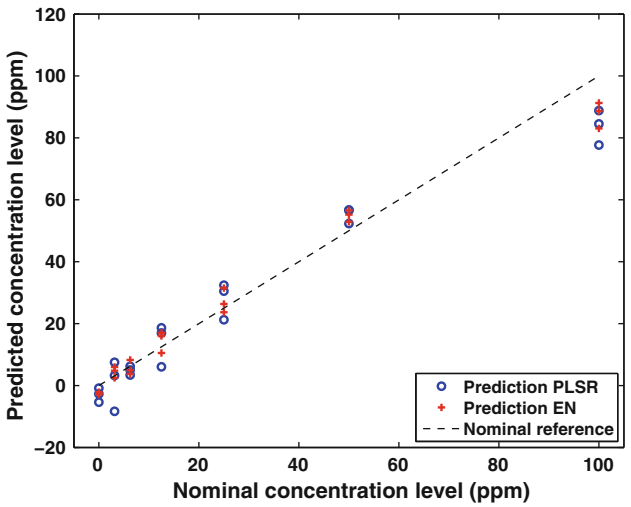

Fig. 18 Prediction of synthetic astaxanthin coating using LOOCV on image samples from the VideometerLab, $n = 70$, using PLSR with three components, and using EN with eight non-zero coefficients

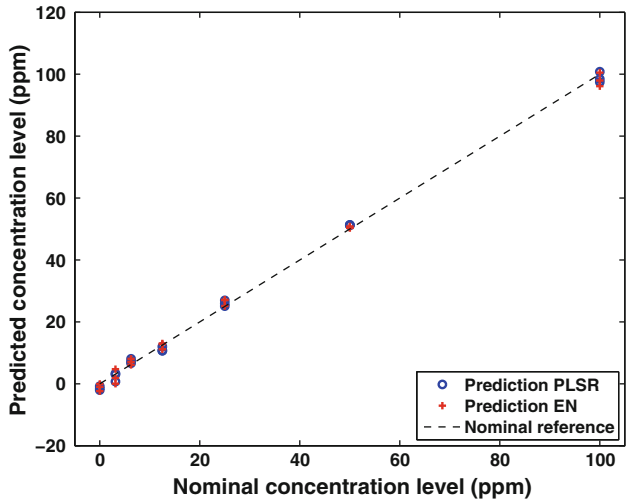

Fig. 19 Prediction of synthetic astaxanthin coating using LOOCV on image samples from the SuperK, $n = 70$, using PLSR with seven components, and using EN with 19 non-zero coefficients

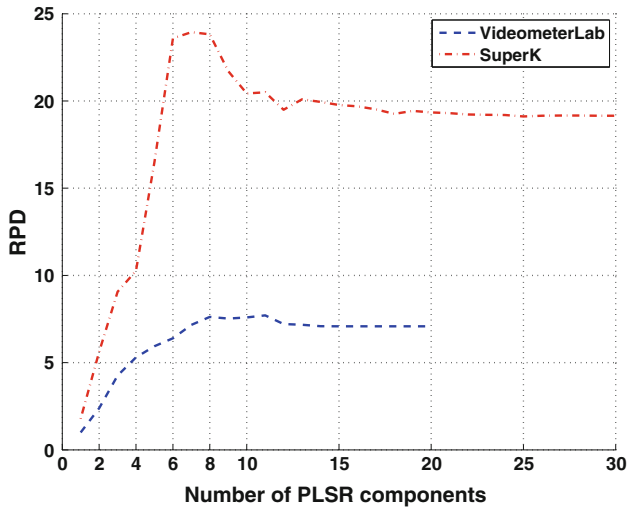

Fig. 20 RPD for different numbers of PLSR components. Using the calibration and validation data set of VideometerLab and SuperK respectively, $n = 70$. The first 30 components are shown for SuperK

For concentration prediction, PLSR and EN models were trained on the data from each instrument and both systems were able to perform results suitable for industrial screening and quality inspection, as the high RPD values between 4 and 24 presented in Table 3 indicate.

PLSR components 1 and 2 include a small amount of all spectral bands and these alone can explain a moderate part of the variance in the data corresponding to the reference values, see Figs. 21 and 22. This could be interpreted such that the overall intensity in all spectral bands explains a part of the synthetic astaxanthin concentration level. However, this is not enough for quality control for the VideometerLab data with respect to RPD value, and just enough for quality control for the SuperK data, as can be seen in Fig. 20. This can be compared to the results in [11] where they found astaxanthin concentration in fish fillets to be strongly dependent on the overall pixel intensity in multispectral images, stated both for PC1 and PLSR component 1.

However, for astaxanthin in pellet coating, it seems as if the concentration level has a more subtle dependency on pixel intensity, since the characteristics appear first in PC2, see Figs. 15 and 16, and in PLSR component 2 and onwards, indicating that the concentration variance is a smaller portion of the data variance. This is common for variance of interest in image analysis [14]. In a previous study of astaxanthin

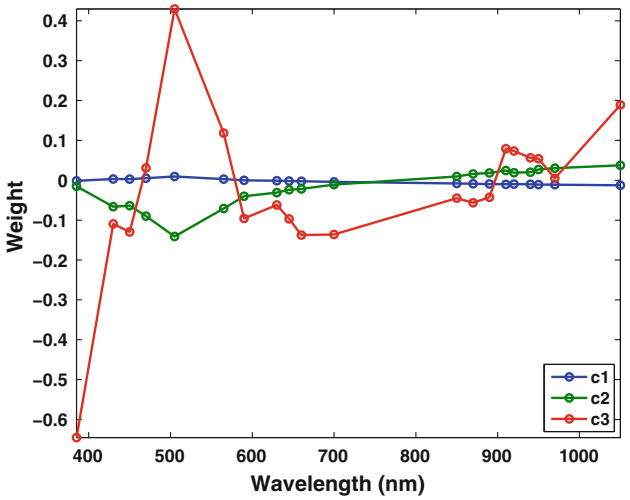

Fig. 21 PLSR components for the prediction of synthetic astaxanthin coating concentration level using LOOCV on image samples from the VideometerLab, $n_c = 49$, $n = 70$. Calculated on the calibration set, showing the three PLSR components

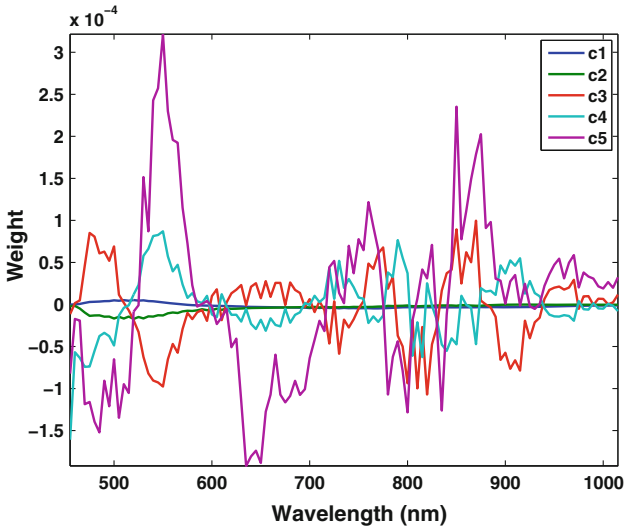

Fig. 22 PLSR components for the prediction of synthetic astaxanthin coating concentration level using LOOCV on image samples from the SuperK, $n_c = 49$, $n = 70$. Calculated on the calibration set, here showing the first five of the total seven PLSR components

coating, PC2 was also found to be of importance for detecting the presence of astaxanthin [20,21]. Furthermore, for astaxanthin coating concentration prediction, PLSR component 2 and onwards showed considerable characteristics for prediction in [23] and component 4 and onwards in [22].

The RPD test shown in Fig. 20 illustrates that more than about 10 PLSR components do not improve the results, which is confirmed by Fig. 17. We therefore conclude that the regression problem for this data has a low complexity since 3 components suffice, though still with a considerable prediction error, and this is considerably improved by a few more components up to about seven or ten components.

The regression models show similar structures between VideometerLab and SuperK, both using PLS regression and EN regression. The PLS regression chose primarily two spectral regions for discriminating between the coating concentrations. The weight of the PLSR components and EN coefficients corresponds well with the variance seen in the spectral difference of the different astaxanthin levels shown in Figs. 13 and 14. The weight of the PLSR coefficients also corresponds well with the primary absorbance peaks found of around 450–505 nm in previous studies of astaxanthin, and both PLSR and EN coefficients correspond well with the secondary absorption peaks of around 500–600 and 870 nm. This also corresponds to the spectroscopy results in Fig. 11.

Since the PLSR component weights are somewhat clustered to the above-mentioned regions, it makes sense to use these regions for making optical filters for an industrial inspection system. The weights of the EN coefficients are clustered in two main parts with opposite signs in the visual range for both modalities, which suggests that the prediction in future could be made using two optical filters. Optical filters can make the equipment cheaper and faster for industrial quality inspection of astaxanthin. The optimum design of such filters can be estimated using sparse methods such as EN or filter-focused methods [19,23,25].

For the EN results it can be seen that the RPD increases with the number of coefficients. While more coefficients in some cases can increase the result, it also enables EN to select different clusters of correlated variables.

To directly compare the performance of the two vision systems, the overlapping 16 bands from the two were used for prediction and the results are presented in Table 3. The VideometerLab data show a prediction error of 13 ppm, while the SuperK have a prediction error of 4 ppm. This is too high for the industry and thus means that many spectral bands are important for a result with high accuracy for this particular prediction problem.

The SEP for the VideometerLab is about 5–8 ppm of synthetic astaxanthin concentration using PLSR and EN, and the SEP for SuperK with three PLSR components is about 3.7 ppm, which means that the error is larger than the smallest level of synthetic astaxanthin; 3.125 ppm. This means that the VideometerLab results are uncertain for this particular level; perhaps this is the limit of the used system. It can be noted that there is an overlap between the smallest astaxanthin concentration levels seen in the PCA plots in Figs. 15 and 16.

Both PLSR and EN perform well for the prediction problem presented. However, the EN prediction model (coefficients) is more interpretable than the PLSR model (components); compare Figs. 21 and 22 with Fig. 24. While PLSR includes all spectral bands in several components in the

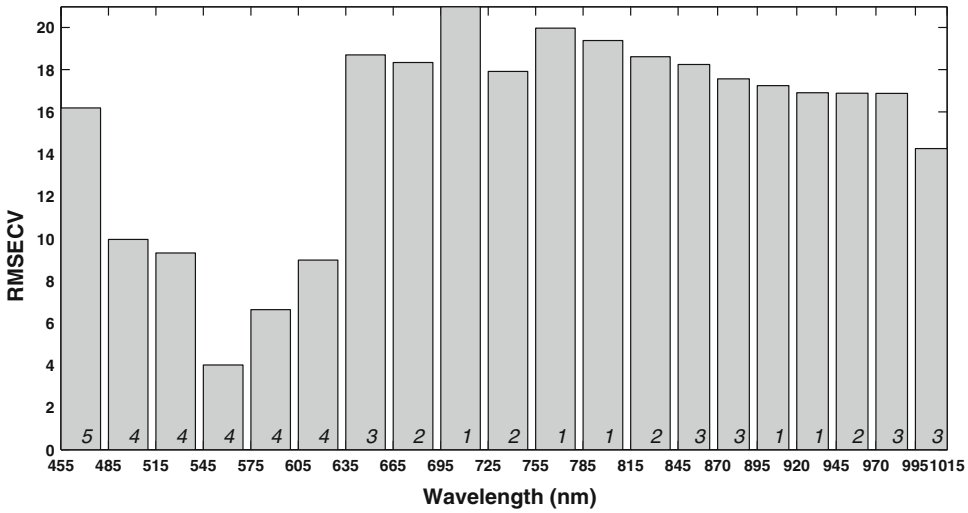

Fig. 23 The iPLS regression calibration using LOOCV on SuperK images illustrating the error (RMSECV) of 20 different regions in the spectrum, $n_c = 49$, $n = 70$. The optimal number of PLSR components chosen for each interval is shown in italics

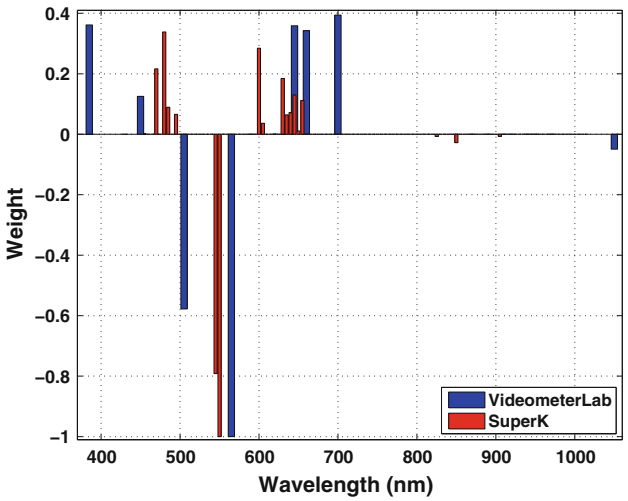

Fig. 24 The coefficients in β_{en} from the prediction model of synthetic astaxanthin coating using EN regression with LOOCV on image samples, $n = 70$, from the VideometerLab using 8 non-zero coefficients, and SuperK using 19 non-zero coefficients. The coefficients are normalised for visualisation

prediction model, though weighted, EN selects just a few spectral bands which make the model easier to interpret and also makes it possible to use a low number of spectral bands.

For testing the generalisation of the PLSR prediction model, a validation scheme here called GFCV was used for leaving out one concentration level at a time for validation. This gives an indication on how the prediction behaves for unseen concentration levels. It is natural to expect a somewhat higher error in this case since the model has not been trained on all concentration levels. However, the generalisation result of GFCV is considered a more honest result than using LOOCV. It is shown that the pooled prediction error of GFCV is much larger than the corresponding error of LOOCV for both systems, see Table 4. The large difference in prediction error implies that the LOOCV scheme in combination with data with many variables compared to the number of samples gives some over-fitting. This can also be assumed from the very optimistic results of both PLSR and EN for the SuperK data with 113 variables, compared to the VideometerLab's 20 variables. The optimistic results using LOOCV on all concentration levels are partly explained by the fact that the study contained only one production batch. It is clear that as the complexity of the reflection increases, the background may have very different contributions from different chemical compounds that all contribute in the surface reflection.

It is therefore concluded that to make a robust and precise prediction model for synthetic astaxanthin coating concentration, it is important to use all target concentration levels, and also that future work examines different production batches.

However, as mentioned previously, HPLC measurements of astaxanthin from fish feed pellets are less accurate than when measuring astaxanthin in oil. A clear difference could be seen in the two measurements of the pellets in Table 2. Therefore, spectral imaging could be a good complement for screening of synthetic astaxanthin coating.

5 Conclusions

A new instrument for hyperspectral imaging, the SuperK setup, based on a spectral broad laser light source, was introduced together with a parallel study with the commercially available multispectral VideometerLab imaging system. We have shown that the new SuperK system can be used for chemical surface inspection using hyperspectral image analysis.

The results show that it is possible to predict the synthetic astaxanthin concentration in the coating well enough for quality control using either multi- or hyperspectral image analysis. Results also show that the SuperK setup performs

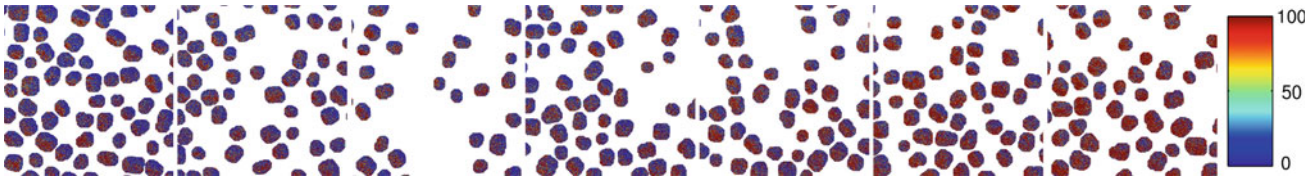

Fig. 25 Projected VideometerLab images using the PLSR model, visualised using the jet colour map where 0 ppm astaxanthin is blue, and 100 ppm astaxanthin is shown as red. The images are taken from the validation set, cropped, clamped and masked. The pellets with different levels of synthetic astaxanthin coating concentration, from left to right: 0, 3.125, 6.25, 12.5, 25, 50, 100 ppm

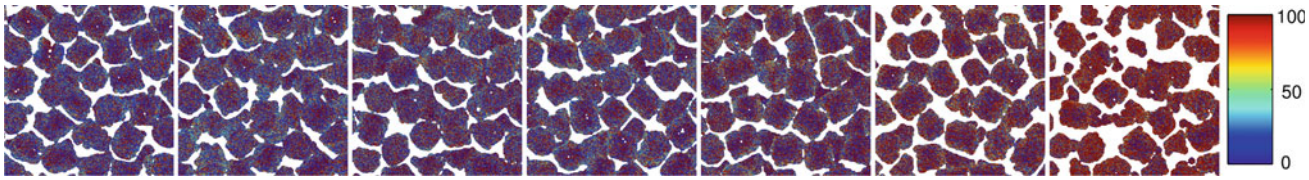

Fig. 26 Projected SuperK images using the PLSR model, visualised using the jet colour map where 0 ppm astaxanthin is blue, and 100 ppm astaxanthin is shown as red. The images are taken from the validation set, cropped, clamped and masked. The pellets with different levels of synthetic astaxanthin coating concentration, from left to right: 0, 3.125, 6.25, 12.5, 25, 50, 100 ppm

with higher accuracy than the VideometerLab for predicting the synthetic astaxanthin concentration in the pellets, while the VideometerLab performs well enough for quality control. The results were obtained by only measuring surface reflections, which in combination with the good results implies that the methods can be used in general for quality inspection of various coating substances using similar coating methods.

In addition, the higher spectral resolution of the SuperK system combined with sparse statistics for analysing the signals made it possible to identify the most significant spectral regions for the particular detection of astaxanthin. This is of interest for a simple and robust commercial system.

Acknowledgments The work presented has received funding from BioMar A/S and the EU under the Seventh Framework Programme FP7/2007-2013 under grant agreement number 214505.10. This work was in part financed by the Centre for Imaging Food Quality project which is funded by the Danish Council for Strategic Research (contract no 09-067039) within the Programme Commission on Health, Food and Welfare. The expert technical assistance of Heidi Olander Petersen is gratefully acknowledged.

References

1. Amarie, S., Förster, U., Gildenhoff, N., Dreuw, A., Wachtveitl, J.: Excited state dynamics of the astaxanthin radical cation. Chem. Phys. **373**(1–2), 8–14 (2010). doi:10.1016/j.chemphys.2010.01.017
2. Baker, R., Pfeiffer, A.M., Schöner, F.J., Smith-Lemmon, L.: Pigmenting efficacy of astaxanthin and canthaxanthin in fresh-water reared atlantic salmon, salmo salar. Anim. Feed Sci. Technol. **99**(1–4), 97–106 (2002)
3. Bligh, E., Dyer, W.: A rapid method of total lipid extraction and purification. Can. J. Biochem. Physiol. **37**(8), 911–917 (1959)
4. Buchwald, M., Jencks, W.P.: Optical properties of astaxanthin solutions and aggregates. Biochemistry **7**(2):834–843 (1968). doi:10.1021/bi00842a042; http://pubs.acs.org/doi/abs/10.1021/bi00842a042
5. Carstensen, J.M., Folm-Hansen, J.: An apparatus and a method of recording an image of an object (2003), EP1051660
6. Clemmensen, L.H., Hansen, M., Ersbøll, B.K.: A comparison of dimension reduction methods with application to multi-spectral images of sand used in concrete. Mach. Vis. Appl. 21:959–968 (2010). doi:10.1007/s00138-009-0193-z
7. Clemmensen, L.H., Hansen, M.E., Frisvad, J.C., Ersbøll, B.K.: A method for comparison of growth media in objective identification of penicillium based on multi-spectral imaging. J. Microbiol. Methods **69**(2), 249–255 (2007). doi:10.1016/j.mimet.2006.12.020
8. Clemmensen, L.K.H., Ersbøll, B.K.: Multispectral recordings and analysis of psoriasis lesions. In: Proceedings of the MICCAI 06—Workshop on Biophotonics Imaging for Diagnostics and Treatment, 9th MICCAI Conference, 6 October 2006 (2006)
9. Dissing, B., Ersbøll, B., Adler-Nissen, J.: New vision technology for multidimensional quality monitoring of food processes. Ph.D. thesis, Technical University of Denmark (DTU) (2011)
10. Dissing, B.S., Carstensen, J.M., Larsen, R.: Multispectral colormapping using penalized least square regression. J. Imaging Sci. Technol. 54(3):0304,011–0304,016 (2010). doi:10.2352/J.ImagingSci.Technol.2010.54.3.030401
11. Dissing, B.S., Nielsen, M.E., Ersbøll, B.K., Frosch, S.: Multispectral imaging for determination of astaxanthin concentration in salmonids. PLoS One **6**(5), Article No.: e19,032 (2011). doi:10.1371/journal.pone.0019032
12. Fernandez-Ahumada, E., Roger, J.M., Palagos, B., Guerrero, J.E., Perez-Marin, D., Garrido-Varo, A.: Multivariate near-infrared reflection spectroscopy strategies for ensuring correct labeling at feed bagging in the animal feed industry. Appl. Spectrosc. **64**(1), 83–91 (2010)
13. Folm-Hansen, J.: On chromatic and geometrical calibration. Ph.D. thesis, Technical University of Denmark (1999)
14. Geladi, P., Isaksson, H., Lindqvist, L., Wold, S., Esbensen, K.: Principal component analysis of multivariate images. Chemom. Intell. Lab. Syst. **5**(3):209–220 (1989); doi:10.1016/0169-7439(89)80049-8. http://www.sciencedirect.com/science/article/B6TFP-44J0P9V-4N/2/6c2915e1439ff9bf8dbe5d19e03a4988
15. Gomez, D.D., Clemmensen, L.H., Ersbøll, B.K., Carstensen, J.M.: Precise acquisition and unsupervised segmentation of

multi-spectral images. Comput. Vis. Image Underst. **106**(2–3), 183–193 (2007). doi:10.1016/j.cviu.2006.06.011
16. Gormley, T.R.: A note on consumer preference of smoked salmon colour. Irish J. Agr. Food Res. **31**(2), 199–202 (1992)
17. Hansen, M.E., Ersbøll, B.K., Carstensen, J.M., Nielsen, A.A.: Estimation of critical parameters in concrete production using multispectral vision technology. Lecture Notes on Computer Science, pp. 1228–1237 (2005)
18. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd edn. Springer, Berlin (2009)
19. Kobayashi, K.I., Nishino, K., Dissing, B.S., Mori, M., Toyota, T., Nakauchi, S.: Design of characteristics of optical filter set for prediction and visualization of fat content in raw beef cuts. In: Proceedings of the Scandinavian Workshop on Imaging Food Quality 2011, Ystad, 27 May 2011, pp. 23–28 (2011)
20. Ljungqvist, M., Ersbøll, B., Frosch, S.: Multivariate image analysis for quality inspection in fish feed production. Ph.D. thesis, Technical University of Denmark (2012)
21. Ljungqvist, M.G., Ersbøll, B.K., Nielsen, M.E., Frosch, S.: Multispectral image analysis for astaxanthin coating classification. J. Imaging Sci. Technol. **56**(020), 403 (2012). http://imaging.org/IST/store/epub.cfm?abstrid=45215
22. Ljungqvist, M.G., Frosch, S., Nielsen, M.E., Ersbøll, B.K.: Multispectral image analysis for robust prediction of astaxanthin coating. Appl. Spectrosc. **67**(7) (2013, to appear)
23. Ljungqvist, M.G., Kobayashi, K.I., Frosch, S., Nielsen, M.E., Ersbøll, B.K., Nakauchi, S.: Near-infrared hyper-spectral image analysis of astaxanthin concentration in fish feed coating. In: Proceedings of the IEEE International Conference on Imaging Systems and, Techniques, pp. 136–141 (2012). doi:10.1109/IST.2012.6295524
24. McConnell, G., Riis, E.: Photonic crystal fibre enables short-wavelength two-photon laser scanning fluorescence microscopy with fura-2. Phys. Med. Biol. **49**(20), 4757 (2004). http://stacks.iop.org/0031-9155/49/i=20/a=007
25. Nakauchi, S., Nishino, K., Yamashita, T.: Selection of optimal combinations of band-pass filters for ice detection by hyperspectral imaging. Opt. Express **20**(2), 986–1000 (2012). doi:10.1364/OE.20.000986
26. Nielsen, O.H.A., Dahl, A.L., Larsen, R., Aanæs, H., Carstensen, J.M., Møller, F., Nielsen, F.D., Thomsen, C.L.: Supercontinuum light sources for hyperspectral subsurface laser scattering: applications for food inspection. Lecture Notes on Computer Science, LNCS, vol. 6688, pp. 327–337 (2011). doi:10.1007/978-3-642-21227-7_31
27. Nielsen, O.H.A., Dahl, A.L., Larsen, R., Møller, F., Nielsen, F.D., Thomsen, C.L., Aanæs, H., Carstensen, J.M.: In depth analysis of food structures. In: Proceedings of the Scandinavian Workshop on Imaging Food Quality 2011, Ystad, 27 May 2011, pp. 29–34 (2011)
28. Nørgaard, L., Saudland, A., Wagner, J., Nielsen, J.P., Munck, L., Engelsen, S.B.: Interval partial least-squares regression (ipls): a comparative chemometric study with an example from near-infrared spectroscopy. Appl. Spectrosc. **54**(3), 413–419 (2000)
29. Ostrander, J., Martinsen, C., Liston, J., McCullough, J.: Sensory testing of pen-reared salmon and trout. J. Food Sci. **41**(2), 386–390 (1976). doi:10.1111/j.1365-2621.1976.tb00626.x
30. Rinnan, A., van den Berg, F., Engelsen, S.B.: Review of the most common pre-processing techniques for near-infrared spectra. TrAC Trends Anal. Chem. **28**(10):1201–1222 (2009). doi:10.1016/j.trac.2009.07.007; http://www.sciencedirect.com/science/article/B6V5H-4WXC219-1/2/9ed550b8d16cca3906099571cf81e32a
31. Sabin, G.P., Breitkreitz, M.C., de Souza, A.M., da Fonseca, P., Calefe, L., Moffa, M., Poppi, R.J.: Analysis of pharmaceutical pellets: an approach using near-infrared chemical imaging. Anal. Chim. Acta **706**(1), 113–119 (2011). doi:10.1016/j.aca.2011.08.029
32. Scherer, N., Jureller, J.E., Birks, T., Wadsworth, W., Russell, P.S.J.: Widely tunable femtosecond pulses from a tapered fiber for ultrafast microscopy and multiphoton applications. In: The Thirteenth International Conference on Ultrafast Phenomena, p. TuE41. Optical Society of America, USA (2002). http://www.opticsinfobase.org/abstract.cfm?URI=UP-2002-TuE41
33. Serra, J.: Image Analysis and Mathematical Morphology. Academic Press, London (1982)
34. Sigurgisladottir, S., Parrish, C., Lall, S., Ackman, R.: Effects of feeding natural tocopherols and astaxanthin on atlantic salmon (salmo salar) fillet quality. Food Res. Int. **27**(1), 23–32 (1994)
35. Simopoulos, A.P.: Omega-3 fatty acids in health and disease and in growth and development. Am. J. Clin. Nutr. **54**, 438–463 (1991)
36. Sjöstrand, K., Clemmensen, L.H., Larsen, R., Ersbøll, B.: Spasm: a matlab toolbox for sparse statistical modeling. J. Stat. Softw. (2013, to appear)
37. Sjöström, M., Wold, S., Lindberg, W., Persson, J.Å., Martens, H.: A multivariate calibration problem in analytical chemistry solved by partial least-squares models in latent variables. Anal. Chim. Acta **150**, 61–70 (1983)
38. Torrisen, O., Hardy, R., Shearer, K.: Pigmentation of salmonids—carotenoid deposition and metabolism. Rev. Aquat. Sci. **1**(2), 209–225 (1989)
39. Williams, P.C., Sobering, D.C.: Comparison of commercial near infrared transmittance and reflectance instruments for analysis of whole grains and seeds. J. Near Infrared Spectrosc. **1**(1), 25–32 (1993)
40. Yuan, J.P., Chen, F.: Identification of astaxanthin isomers in haematococcus lacustris by HPLC-photodiode array detection. Biotechnol. Tech. **11**(7), 455–459 (1997)
41. Zou, H., Hastie, T.: Regularization and variable selection via the elastic net. J. R. Stat. Soc. Ser. B Stat. Methodol. **67**, 301–320 (2005)

Author Biographies

Martin Georg Ljungqvist received his M.Sc. Engineering degree from Linköping University in 2005 and received his Ph.D. in multispectral image analysis in 2012 from DTU Informatics, Department of Applied Mathematics and Computer Science at the Technical University of Denmark. He has also worked with programming of signal processing and image analysis for embedded systems.

Otto Højager Attermann Nielsen received his M.Sc. degree from Aarhus University in 2009 and since 2010 he is a Ph.D. Student in optimising vision systems at the Department of Applied Mathematics and Computer Science at the Technical University of Denmark.

Stina Frosch received her M.Sc. degree in food science from The Royal Veterinarian and Agricultural University in 2000 and her Ph.D. in traceability and multivariate data analysis in 2006 from the Technical University of Denmark. She is currently working as an associate professor in food science at the National Food Institute, Division of Industrial Food Research at the Technical University of Denmark.

Michael Engelbrecht Nielsen got his M.Sc. degree in biology in 1994 and his Ph.D. in fish immunology in 1998 from Copenhagen University. He has worked as an associate professor at The Royal Veterinarian and Agricultural University until 2007 and is currently working as an associate professor in biological quality at the National Food Institute, Division of Industrial Food Research at the Technical University of Denmark. He is also a member of several EU projects within aquaculture.

Line Harder Clemmensen received her M.Sc. in applied mathematics in 2006 and received her Ph.D. from the Technical University of Denmark in 2009. Currently, she is an assistant professor in the Statistics and

Data Analysis section at the department of Informatics and Mathematical Modelling, at the Technical University of Denmark. Her research interests include image analysis, over-complete problems, dimension reduction, and statistical pattern recognition.

Bjarne Kjær Ersbøll received his M.Sc. Engineering and Ph.D. degrees in statistics and image analysis from the Technical University of Denmark in 1983 and 1990, respectively. He is currently a professor at the department of Informatics and Mathematical Modelling, the Technical University of Denmark, where he is the head of the Statistics and Data Analysis section. His research interests include applied multivariate statistical analysis of multispectral digital imagery. He has worked on related problems from many domains: remote sensing, industry, and medical/biological. He has also organised or co-organised many international and national conferences, and workshops on these topics.