ELSEVIER

Journal of Food Composition and Analysis 20 (2007) 73–79

JOURNAL OF FOOD COMPOSITION AND ANALYSIS

www.elsevier.com/locate/jfca

Original Article

Rapid estimation of fat content in salmon fillets by colour image analysis

Lars Helge Stien[a,*], Anders Kiessling[b,c], Fredrik Manne[d]

[a]*Department of Biology, University of Bergen, PO Box 7800, No-5020 Bergen, Norway*
[b]*Institute of Marine Research, Matre Research Station, No-5984 Matredal, Norway*
[c]*Department of Animal and Aquaculture Sciences, Norwegian University of Life Sciences, PO Box 5003, No-1432 Ås, Norway*
[d]*Department of Informatics, University of Bergen. Thormøhlensgt, 55. No-5020 Bergen, Norway*

Received 12 July 2005; received in revised form 18 July 2006; accepted 18 July 2006

Abstract

The aim of this study was to develop a simple method for the automatic estimation of fat content in salmon fillets by means of image analysis. Salmon fillets have a zebra-like appearance, consisting of white myocommata stripes divided by red-coloured muscle tissue. The stripes contain a high proportion of lipid, and it has previously been shown that the proportion of myocommata in a fillet correlates with its fat content. A possible measurement method for fillet fat content might therefore be to use image analysis to determine the area of the white stripes visible on the fillets surface compared to the total area of the fillet. Fifteen salmon fillets were sampled from an assembly line at a local fish-processing plant, photographed and analysed for lipid. The results obtained by the image analysis were compared with those from chemical analysis, and showed a good correlation ($r = 0.84$). Although the sample size was relatively small, the correlation was high enough to suggest that the method, after development and improvement, could prove useful as a simple tool in salmon aquaculture research and product processing.

Keywords: Lipid; Fat; Salmon; Fillet; Image analysis

1. Introduction

Recent advances in the area of computers and image processing have created new ways to monitor quality in the food industry (Brosnan and Sun, 2004). The salmon industry and research community are in need of fast, reliable and economical methods of quality assessment. Image analysis is capable of providing a wide range of information about a product from a single image in a fraction of a second, making it possible to analyse products as they pass on a conveyor belt (Polder et al., 2004; Storbeck and Dan, 2001). Image analysis is also non-destructive and therefore ideal for monitoring the quality of fish products. Kohler et al. (2002), for instance, developed a method for sorting quality classes of cod fillets. Sorting salmon fillets by quality could involve quantifying characteristics such as fat content, shape and colour. Fat content is one of the most important characteristics of salmon fillet quality, and may vary widely from fillet to fillet; a study carried out in 1995 found that the fat content of Norwegian salmon fillets ranged from 11% to 19% (Fjellanger et al., 2000). Lipid content also influences other quality characteristics such as gaping (Shearer, 2000) and colour (Rørå et al., 1998). It is therefore important to have rapid, reliable methods of lipid measurement.

There exist several methods for estimating the lipid content of fish: chemical analysis, the Torry fat meter, computerized tomography (CT) and near-infra-red (NIR) spectrophotometry and calibration (Fjellanger et al., 2000). Chemical analysis is laborious, and both CT and NIR require expensive equipment. The image analysis method described in this paper is non-destructive, demands a minimum of labour and the necessary equipment is relatively inexpensive. The method is also very fast,

*Corresponding author. Tel.: +47 55584608; fax: +47 55589611.
E-mail address: lars.stien@bio.uib.no (L.H. Stien).

0889-1575/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.jfca.2006.07.007

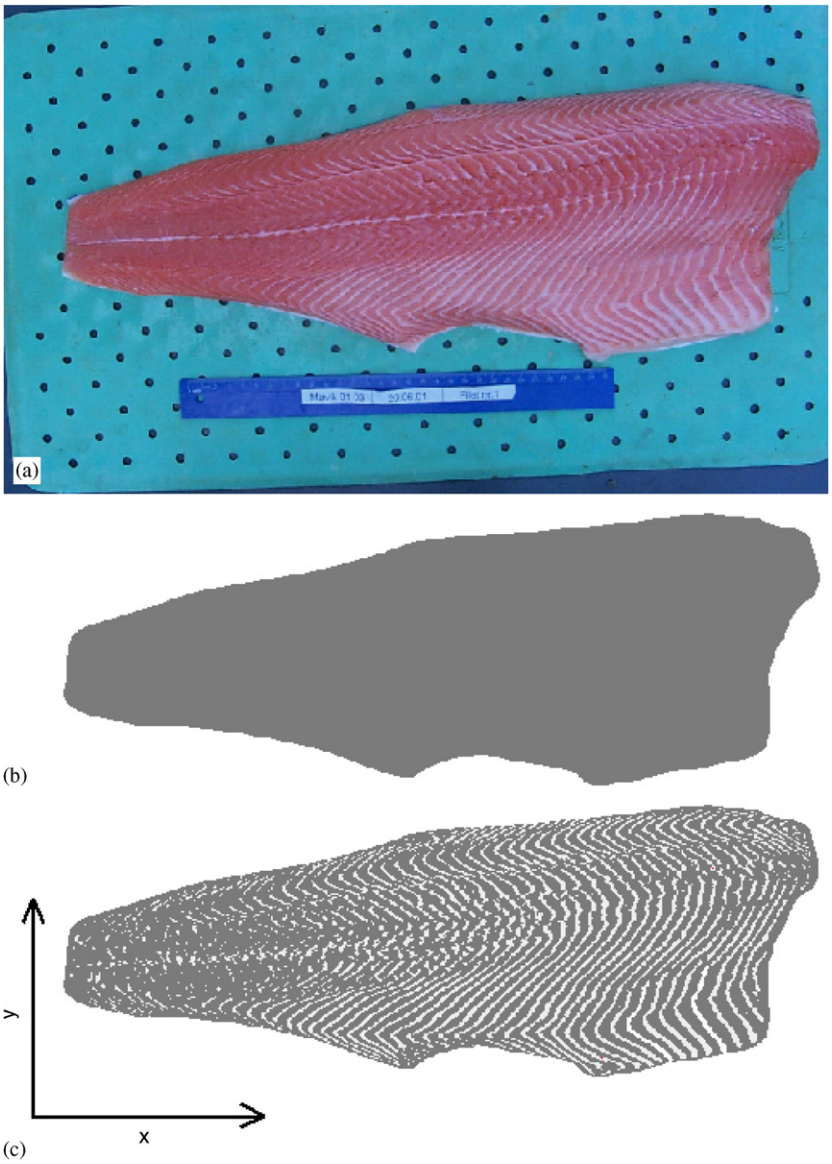

Fig. 1. Different stages of the image analysis. (a) original image; (b) the fillet region has been identified by Eq. (1) and (c) discrimination of the fat stripes by Eq. (2) with parameters set to $T = 5$ and $m = 21$. The directions of the x- and the y-axis are also shown.

needing less than 1 s to compute on a state-of-the-art personal computer. Another advantage is that the images used by the method to estimate fat content probably also can be used to measure other quality traits such as colour and shape. This is not the focus of our present work, but will be presented under separate cover.

A large proportion of the lipid in a salmon fillet is congregated in white stripes of connective tissue (myocommata) segmenting the red-coloured muscle tissue in vertical blocks and giving the fillet a zebra-like appearance (Fig. 1a). The region between the stripes consists mainly of muscle tissue, and is referred to here as meat or muscle. The muscle also contains fat, but to a much lesser extent than the white stripes. The size of the white stripes correlates well with the fat content of the fillet. This has been used to estimate the fat content from CT scans of whole salmon (Rye and Gjerde, 1991; Bjerkeng et al., 1997). In these studies the fat content of whole fish was estimated as the fraction of the area of fat deposits versus the total CT scan area of the fish. This makes it reasonable to assume that the area of visible stripes on a salmon fillet divided by the total area of the fillet will also correlate with the fat content of the fillet. A semi-automatic image analysis technique based on this assumption has already been tried (Borderías et al., 1999), but it obtained a relatively low correlation coefficient with chemically measured fat content; $r = 0.41$. A much higher correlation was achieved by Marty-Mahé et al. (2004), who measured fat content from images of brown trout (*Salmo trutta* L.) cutlets; $r = 0.76$. Measuring the fat content from images of cutlets is different from that of fillets due to the different organization of the connective tissue. However, this later study of cutlets suggests that higher correlations than $r = 0.41$ could also be obtained for fillets if the appropriate

image analysis method were used. This study presents a fully automatic image analysis method for estimating the fat content of salmon fillets. The method is relatively simple, and only basic programming skills are necessary to duplicate the method.

2. Material and methods

Fifteen fully trimmed fillets of Atlantic salmon (*Salmo salar* L.) were collected from a commercial producer (Lerøy, Bergen, Norway) and stored on ice for 24 h before photography. The fillets weighed from 1.5 to 2 kg, representing a standard size offered to the market. The fillets were placed skin down on a horizontal green plastic base and photographed directly from above using a digital camera (Kodak DC260). The images were taken outdoors in shadow, avoiding direct sunlight, which provided images without specular reflection (Ray, 1999) (Fig. 1a). The camera provided Red–Green–Blue (RGB) images in the JPEG format (Miano, 1999) at a resolution of 1136×1024 pixels. Each pixel in an RGB image is represented by three colour values R, G and B spanning from minimum 0 to maximum 255. In the following, the colour values of each pixel (x,y), where $x \in \{1,2,\ldots,1136\}$ and $y \in \{1,2,\ldots,1024\}$, are donated as $R(x,y)$, $G(x,y)$ and $B(x,y)$, respectively. An object of known area, placed in one of the images, showed that 1 cm^2 in the scene corresponded to an image area of 23×23 pixels. For the subsequent chemical fat analysis each fillet was minced and total lipid determined gravimetrically from a 5 g subsample after extraction with ethyl acetate (Losnegard et al., 1979). The chemical fat analysis was performed in duplicate.

The proposed image analysis method consists of two main steps: discrimination of the fillet from the background, followed by discrimination of the lipid stripes within the fillet. The digital images contain two main regions, the green base or background, and the respective fillet (Fig. 1a). The white stripes, although appearing white to the eye as compared to the adjacent dark red muscle, are in fact light red in the images. The fillet in each image can therefore be separated from the background by stating that all red pixels in the image are fillet, while all others are background. In other words, all pixels (x,y) in the image with an R-colour value higher than both its G- and B-colour value are classified as fillet

$$\alpha(x,y) = \begin{cases} 1 \Leftrightarrow R(x,y) > G(x,y,) \wedge R(x,y) > B(x,y) \\ 0 \quad \text{otherwise,} \end{cases} \tag{1}$$

where $\alpha(x,y)$ equals 1 indicates fillet and $\alpha(x,y)$ equals 0 indicates background (Fig. 1b).

The next step was to discriminate the lipid stripes from the other parts of the fillet. The white stripes are difficult to distinguish in the red R-colour layer (Fig. 2a). They have, however, clearly higher (in effect lighter) values compared to adjacent muscle in the G- and B-colour layers (Fig. 2b and c). Simply using a global threshold as in the study by Borderías et al. (1999) resulted in large areas of false segmentation for all possible threshold values in both the G- and B-colour layer. The method suggested for cutlets by Marty-Mahé et al. (2004) (K_{mean} and Profile) was also based on a global threshold. In the images, the muscle in the abdomen typically has the same colour as the fat stripes in other parts of the fillet. This is probably caused both by non-uniform lighting on the scene and differences in pigmentation between different parts of the fillet. One solution to this problem is to discriminate the stripes from the muscle based on local information in the image, for instance the average colour value in a defined neighbourhood around each pixel. The proposed method first smoothes the G-colour layer with a 3×3-median filter (Gonzalez and Woods, 1992) to remove noise. Adaptive thresholding, utilizing the average colour value in an $m \times 1$-horizontal neighbourhood to determine the local threshold, is then used to discriminate the fat stripes from the muscle.

$$\beta(x,y) = \begin{cases} 0 \Leftrightarrow \alpha(x,y) = 0 \\ 1 \Leftrightarrow \alpha(x,y) = 1 \;\&\; G(x,y) \leqslant \frac{1}{N} \sum_{\mu=x-(m-1)/2}^{x+(m-1)/2} G(x+\mu,y) \\ \qquad \times \alpha(x+\mu,y) - T \\ 2 \text{ otherwise} \end{cases}$$

$$N = \sum_{\mu=x-(m-1)/2}^{x+(m-1)/2} \alpha(x+\mu,y), \tag{2}$$

where N is the number of pixels representing fillet in the respective $m \times 1$-neighbourhood. In other words, all pixels (x,y) in the fillet-region, i.e., $\alpha(x,y) = 1$ (Eq. (1)), with a G-value higher than the average G-value in the respective $m \times 1$-neighbourhood plus a constant T were set to represent lipid, $\beta(x,y) = 2$. An $m \times 1$-neighbourhood was chosen because of the close to vertical orientation of the fat stripes (Fig. 1a); a horizontal $m \times 1$-neighbourhood centred any place on the fillet region is likely to contain both pixels representing stripes and muscle. Using an $m \times 1$-neighbourhood is also extremely efficient. For each horizontal line in the image, Eq. (2) is performed on subsequent pixels moving from left to right. Except for the first pixel in each line, the sums in Eq. (2) can be calculated efficiently from the previous step by subtracting the value left of the neighbourhood $(x-(m-1)/2-1,y)$ and adding the new value now included in the neighbourhood $(x+(m-1)/2,y)$. The image analysis method was tested with four different sizes of the $m \times 1$-neighbourhood ($m = 5$, 11, 21 and 31) and four different T values ($T = 0$, 5, 10 and 15). Estimated fat content by image analysis is calculated by the proposed method as

$$\text{IA lipid content (\%)} = 100 \cdot \frac{\sum(\beta = 2)}{\sum(\beta \geqslant 1)}. \tag{3}$$

In the experiment performed by Borderías et al. (1999), only the tail half portion of the fillet was photographed. In order

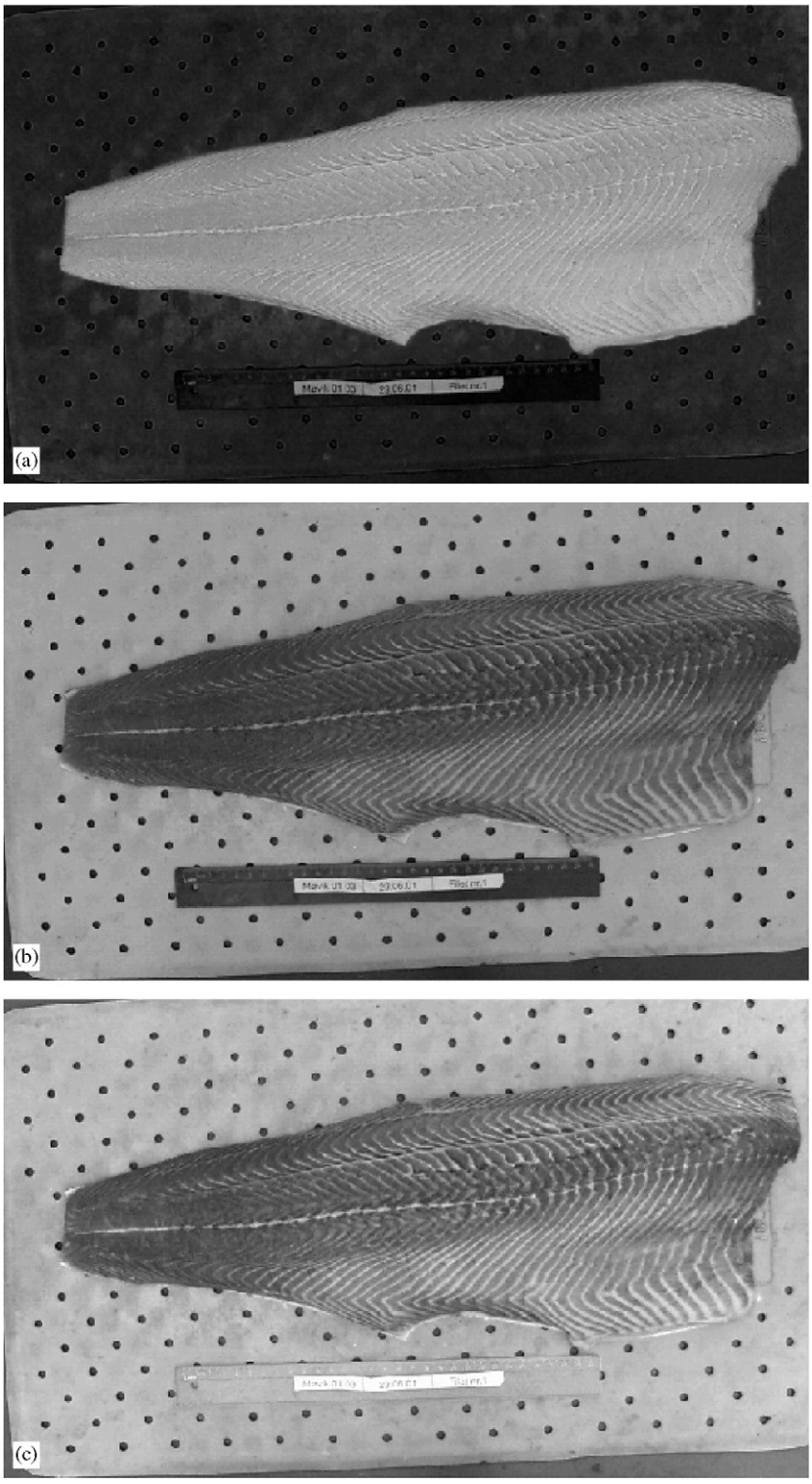

Fig. 2. The intensity images of the (a) (R)ed-, (b) (G)reen- and (c) (B)lue colour layers.

to mimic this experiment the proposed image analysis method was also tried on only the posterior part of the fillet. This was done by first finding the maximum and minimum x-values in the α-matrix equal to 1 and then setting all elements in α with x-coordinate > minimum x + (maximum x − minimum x)/2–0. In other words, the front half portion of the fillet region was set to 0 in α before applying Eq. (2).

Finally, the relationship between chemical and automated image analysis was tested by the Pearson correlation coefficient (Johnson and Bhattacharyya, 2000) and regression lines calculated using the SAS software package (SAS Institute Inc., Cary, North Carolina, USA) for Windows, version 8e (proc corr and proc reg). A P-value <0.05 was considered significant. All data were tested for

normality by a normal probability plot (proc univariate, SAS).

3. Results and discussion

Eq. (2) includes two parameters: m and T. The correlation of estimated lipid content by the image analysis method with chemically measured lipid content is dependent on both these parameters (Table 1). The highest correlation ($r = 0.83$, Figs. 1c and 3a) was achieved when m and T were set to 21 and 5, respectively. Spearman rank coefficient (Johnson and Bhattacharyya, 2000) measured a correlation of $r = 0.92$, indicating that there is an even stronger non-linear relationship between the two variables. Altering the T- and/or m-values from these ideal settings resulted in lower correlation values (Table 1).

The ideal situation for the adaptive thresholding method (Eq. (2)) is when the fat locally has clearly distinct G-values from the meat. In this case the average value of the elements in the $m \times 1$-neighbourhood can be used as a threshold between the two regions (Fig. 4a). However, this ideal, or model, situation is not realistic. Instead there is a sliding transition from muscle to lipid stripe (Fig. 4b and d). If the relative amount of lipid stripe is small compared to the relative amount of muscle, the muscle will dominate the average, pulling it down, and as a result the segmented lipid stripes will become overly wide. Moreover, using only the average value as a threshold will lead to segmentation into two regions even when the neighbourhood contains only pixels representing muscle (Fig. 4c). All in all, this explains the negative correlations in the first column of Table 1. A constant T was therefore added to the average value in Eq. (2) to avoid oversegmentation of lipid stripes. However, this constant must not be set too high, as this would lead to undersegmentation, especially in areas where the lipid stripes are wide (Fig. 4d).

Fig. 4 also shows the effect on the same stripe for a large neighbourhood (Fig. 4b) and a small neighbourhood (Fig. 4d). Selecting an m that is too small will lead to undersegmentation when the neighbourhood is centred on the stripe (Fig. 4d), and oversegmentation when centred on muscle (Fig. 4c). This explains the relatively low correlation for the smallest m tested ($m = 5$, Table 1). If the mask is too large, variations in muscle colour and lighting over the fillet may lead to invalid segmentations.

A high correlation ($r = 0.80$) was also obtained when only the tail-half portion of the fillet was used in the image analysis ($m = 21$, $T = 5$). Comparing the two plots in Fig. 3, there is little reason to suggest that using the whole fillet is superior to only using the posterior. This indicates that one can get a good estimate of the lipid contents by using only the tail half portion of the fillet. This is an

Table 1
Correlations between the results from the chemical analysis and the proposed image analysis method on whole filet at different parameter settings (Eq. (2))

T m	0	5	10	15
5	—	$r = 0.55$ $P = 0.03$	$r = 0.50$ $P = 0.06$	—
11	$r = -0.60$ $P = 0.02$	$r = 0.79$ $P < .0.01$	$r = 0.68$ $P < 0.01$	$r = 0.58$ $P = 0.02$
21	$r = -0.57$ $P = 0.03$	$r = 0.83$ $P < 0.01$	$r = 0.82$ $P < 0.01$	$r = 0.77$ $P < 0.01$
31	—	$r = 0.75$ $P < 0.01$	$r = 0.79$ $P < 0.01$	$r = 0.78$ $P < 0.01$

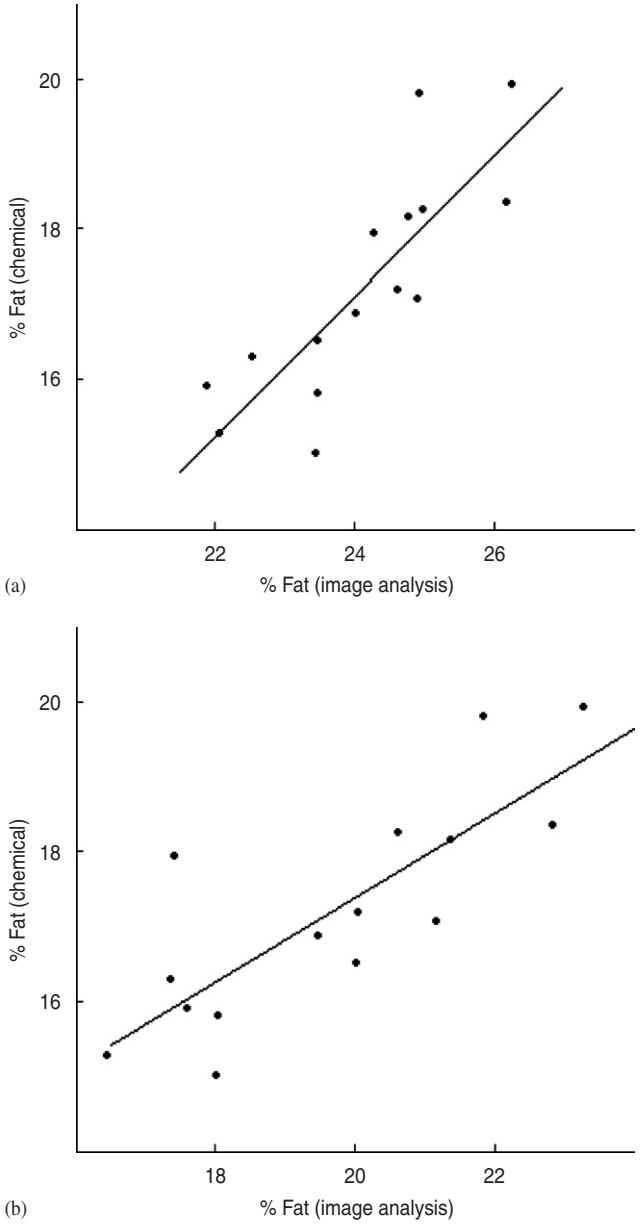

Fig. 3. Chemical measurement of fat content compared to image analysis (Eq. (3)). Parameters in Eq. (2) set to $T = 5$ and $m = 21$. (a) Analysis of whole fillet, regression line: $Y = 0.936X$ ($P < 0.001$)-5.381($P = 0.231$), $R^2 = 0.68$. (b) Analysis on the tail half portion of the fillet. $Y = 0.566X$($P < 0.001$)$+ 6.061$ ($P = 0.019$), $R^2 = 0.65$. Where X is fat content measured by image analysis and Y is fat content measured chemically.

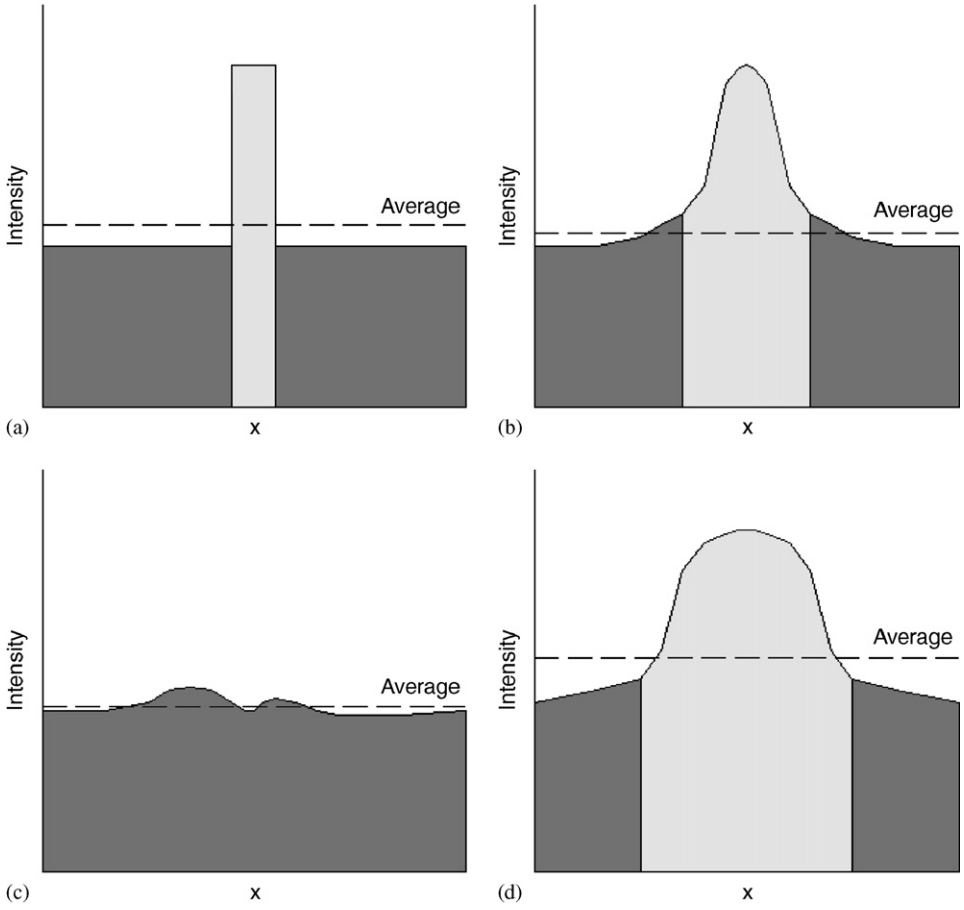

Fig. 4. Schematic representation of the intensity values of the G-colour layer in a $m \times 1$-neighbourhood (Eq. (2)). (a) Ideal situation: the average intensity value divides lipid from muscle. (b) Thin lipid stripe: the average value is dominated by muscle pulling it down. (c) Muscle alone: the average value is lower than many muscle values. (d) Thick lipid stripe: the average value is dominated by lipid.

advantage, as after initial filleting at the processing plant large parts of the peritoneum may be left on the fillet, covering much of the fillet and obscuring lipid stripes in the abdomen. This is why fully trimmed fillets (with all traces of peritoneum removed) were selected for this study. The peritoneum extends also into the tail-half portion of the fillet, but only to a relatively limited degree. Estimation of lipid content by image analysis can therefore probably be performed regardless of the degree of trimming.

When evaluating the correlation between the image analysis and the chemical analysis of lipid content, it is important to remember that a correlation of $r = 1$ is not to be expected. In the first place, the lipid content of both the muscle and connective tissue relative to each other may vary. Secondly, the chemical analysis has a confidence interval of $[-1, 1]$ with 99% probability. Thirdly, the fillets covered a relatively small range of lipid contents, namely 15–20%. All in all, this indicates that the image analysis method correlated well with the results of chemical analysis, and can be used to distinguish between fillets even when there are only relative small differences in the lipid content. However, it is a problem that the method has only been tested on a small number of fillets from a single sampling. The method has therefore not been tested on the whole range of lipid contents and fillet sizes that may turn up on a processing line. Small fillets, for instance, will probably need a smaller m in Eq. (2). A future scenario might involve first calculating the size of the fillets and then automatically choosing parameter settings on this basis.

The basic requirement of the presented method is that a large part of the fish fillet's fat is congregated in white stripes divided by darker muscle tissue. This is the case in most salmonids depositing carotenoids in the somatic muscle. However, fish not displaying such distinct visible differences between fat and muscle require other approaches. Selective use of specific wavelengths, magnetic resonance imaging, infra-red imaging (IR) and computer tomography imaging (CT scanning) can offer possible solutions. Such approaches are most interesting for future work, but were not part of the present study.

4. Conclusions

Image analysis is a promising tool for estimating the fat content not only of salmonid cutlets (Marty-Mahé et al., 2004) but also of salmonid fillets. The method does not require sophisticated equipment or advanced image analysis programs. The image acquisition was performed with a

standard consumer camera, and the image analysis utilized relatively simple mathematical procedures performed on the image matrix. Further testing on larger samples will be necessary before the method can be accepted as a general method for measuring lipid in Salmonid fillets.

Acknowledgements

The Research Council of Norway financially supported the project through project grants nos. 153178/120. The practical part of the project was carried out at Matre Aquaculture Research Station, and we are grateful to the staff for their technical assistance.

References

Bjerkeng, B., Refstie, S., Fjalestad, K.T., Storebakken, T., Rødbotten, M., Roem, A., 1997. Quality parameters of the flesh of Atlantic salmon (*Salmo salar*) as affected by dietary fat content and full-fat soybean meal as a partial substitute for fish meal in the diet. Aquaculture 157, 297–309.

Borderías, A.J., Gómez-Guillén, M.C., Hurtado, O., Montero, P., 1999. Use of image analysis to determine fat and connective tissue in salmon muscle. European Food Research and Technology 209, 104–107.

Brosnan, T., Sun, D-W., 2004. Improving quality inspection of food products by computer vision – a review. Journal of Food Engineering 61, 3–16.

Fjellanger, K., Obach, A., Rosenlund, G., 2000. Proximate analysis of fish with special emphasis on fat. In: Kestin, S.C., Warris, P.D. (Eds.), Farmed Fish Quality. Fishing News Books Ltd., Blackwell Science, Oxford, UK, pp. 307–317.

Gonzalez, R.C., Woods, R.E., 1992. 4.3.2 Smoothing filters. In: Gonzalez, R.C., Woods, R.E. (Eds.), Digital Image Processing. Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, USA, pp. 191–195.

Johnson, R.A., Bhattacharyya, G.K., 2000. Statistics: Principles and Methods. In: Johnson, R.A., Bhattacharyya, G.K. (Eds.). Wiley, New York, USA.

Kohler, A., Skaga, A., Hjelme, G., Skarpeid, H.J., 2002. Sorting salted cod fillets by computer vision: a pilot study. Computers and Electronics in Agriculture 36, 3–16.

Losnegard, N., Bøe, B., Larsen, T., 1979. Undersøkelse av ekstraksjonsmidler for bestemmelse av fett. Fiskeridirektoratet, Bergen, Method no. 1/79 (in Norwegian).

Marty-Mahé, P., Loisel, P., Fauconneau, B., Haffray, P., Brossard, D., Davenel, A., 2004. Quality traits of brown trouts (Salmo trutta) cutlets described by automated color image analysis. Aquaculture 232, 225–240.

Miano, J., 1999. Compressed Image File Formats. In: Miano, J. (Ed.). Addison Wesley Publishing Company, Inc., Reading, Massachusetts, USA, pp. 35–76.

Polder, G., van der Heijden, G.W.A.M., van der Voet, H., Young, I.T., 2004. Measuring surface distribution of carotenes and chlorophyll in ripening tomatoes using imaging spectrometry. Postharvest Biology and Technology 34, 117–129.

Ray, S.F., 1999. Scientific photography and applied imaging. In: Ray, S.F. (Ed.). Focal Press, Oxford, UK, pp. 84–86.

Rye, M., Gjerde, B., 1991. Prediction of carcass composition in Atlantic salmon by computerized tomography. Aquaculture 99, 35–48.

Rørå, A.M.B., Kvåle, A., Mørkøre, T., Rørvik, K.-A., Steien, S.H., Thomassen, M.S., 1998. Process yield, colour and sensory quality of smoked Atlantic salmon (*Salmo salar*) in relation to raw material characteristics. Food Research International 31, 601–609.

Shearer, K.D., 2000. The effect of diet composition and feeding regime on proximate composition of farmed fishes. In: Kestin, S.C., Warris, P.D. (Eds.), Farmed Fish Quality. Fishing News Books Ltd., Blackwell Science, Oxford, UK, pp. 31–41.

Storbeck, F., Daan, B., 2001. Fish species recognition using computer vision and a neural network. Fisheries Research 51, 11–15.