Available online at www.sciencedirect.com
ScienceDirect
Aquaculture 261 (2006) 695–705
Aquaculture
www.elsevier.com/locate/aqua-online

Automated image analysis as a tool to quantify the colour and composition of rainbow trout (*Oncorhynchus mykiss* W.) cutlets

Lars Helge Stien [a,b,*], Fredrik Manne [c], Kari Ruohonene [d], Antti Kause [e], Krisna Rungruangsak-Torrissen [b], Anders Kiessling [b,1]

[a] *Department of Biology, University of Bergen. P.O. Box 7800, NO-5020 Bergen, Norway*
[b] *Institute of Marine Research-Matre, NO-5984 Matredal, Norway*
[c] *Department of Informatics, University of Bergen. Thormøhlensgt. 55. NO-5020 Bergen, Norway*
[d] *Finnish Game and Fisheries Research Institute, Turku Game and Fisheries Research, Itäinen Pitkäkatu 3, 20520 Turku, Finland*
[e] *MTT Agrifood Research Finland, Animal Production Research, Animal Breeding, FIN-31600 Jokioinen, Finland*

Received 15 December 2005; received in revised form 3 August 2006; accepted 3 August 2006

Abstract

The goal of this paper is to propose and evaluate automated image analysis methods for describing muscle cutlets in rainbow trout. The proposed automated image analysis methods were tested on a total of 983 scanned images of trout cutlets, and included quality traits such as fat percentage, flesh colour and the size of morphologically distinguishable subparts of the cutlet. A sub-sample of 50 images was randomly selected for manual segmentation of the cutlet, the dorsal fat depot and the red muscle and regions. The identification of these regions by manual and automatic image analysis correlated strongly (r=0.97, r=0.95 and r=0.91, respectively). The estimated fat percentage obtained from image analysis, based on the area of visible fat and the colour of the cutlet flesh, correlated well with chemical fat percentage measured by mid-infrared transmission spectroscopy (MIT) (r=0.78). The automated image analysis methods are therefore a reliable means of predicting the fat percentage of trout cutlets. Principal component analysis (PCA) loading plots were used to identify subsets of variables from the image analysis of special significance for further studies; cutlet area, dorsal fat depot area, red muscle area, back height, cutlet width, and width of left and right abdomen wall were among the variables selected. PCA loading plots of different colour variables indicated that simple statistical coefficients such as percentiles and mean values can be used to quantify different aspects of flesh colour. In conclusion, the methods presented here provide a powerful toolbox for describing important morphological structures and quality traits of trout cutlets.

Keywords: Image analysis; Rainbow trout; Cutlet; Quality traits

1. Introduction

This study forms part of a larger study that focuses on the genetic components of protein and growth efficiency in salmonids and their relation to important quality traits. The project requires the analysis of a large number

* Corresponding author. Postal address: University of Bergen, Department of Biology, P.O. Box 7800, NO-5020 Bergen, Norway. Tel.: +47 55584608.
E-mail address: lars.stien@bio.uib.no (L. H. Stien).
[1] Department of Animal and Aquaculture Sciences, Agricultural University of Norway, P.O. Box 5003, NO-1432 Ås, Norway (present address).

0044-8486/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.aquaculture.2006.08.009

of fish for quality traits. Automated image analysis, also called machine vision, is an efficient and non-invasive tool for measuring quality traits in fish and other food products (Brosnan and Sun, 2004). It is also widely used in fish research. Borderías et al. (1999) used image analysis to determine the fat percentage of Atlantic salmon fillets (*Salmo salar* L.). The images were made in ultraviolet light in order to enhance the visibility of the fat stripes, and the subsequent image analysis consisted of a Laplace transformation to increase image contrast followed by manual segmentation of the image into shades of grey. The major obstacle to the use of this method in a large-scale study is that the image segmentation step is not fully automatic, while it also resulted in a relatively low correlation with chemically measured fat percentage ($r=0.42$). Rønsholdt et al. (2000) used image analysis of rainbow trout (Oncorhynchus mykiss W.) cutlets scanned on a flatbed scanner as a potential tool to quantify the area of the cutlet, the area of the fat stripes and the area of the muscle. The cutlet was segmented from the background by stating a constant threshold value in the blue colour layer of the RGB colour model. The fat stripes were then segmented from the muscle in a similar fashion, but in the green colour layer. However, this image analysis method was not compared with any established fat analysis method. The number of fish was also relatively small ($n=20$). Marty-Mahé et al. (2004) constructed an elaborate light tent for imaging of brown trout (*Salmo trutta* L.) cutlets. Images of 48 cutlets were made using a digital camera and image analysis methods were developed to quantify the colour, fat stripes, myomera and peripheral surfaces. The highest correlation between the relative area of the fat stripes and fat percentage measured by Soxhlet and NMR ($r=0.72$) was obtained by dividing the image pixels into four different classes (background, muscle, fat stripes and other) using the k-mean thresholding method, followed by a procedure called the Profile method that enhanced the final segmentation of the fat stripes. In addition to these studies on the segmentation of fat stripes from salmonids, there have been several studies on distinguishing fat from red meat by image analysis. Ballerini et al. (2000, 2001), for instance, used NMR images and an automatic segmentation technique called Fuzzy thresholding to quantify intramuscular fat percentage in beef.

The goal of the present study was to develop image analysis methods for quantifying the size and colour of morphological structures of trout cutlets such as fat stripes, red muscle, white muscle, and dorsal fat depot. Scanning is an easy-to-use method of obtaining images. No elaborate lighting scheme, in the form of light boxes or light tents, is needed. We therefore believe that image analysis methods developed for scanned images of salmonid cutlets have the potential to be used by the research community.

2. Materials and methods

2.1. Fish material and sampling

A detailed description of fish management and experimental design is given in Kause et al. (2006). In brief, three-year-old rainbow trout ($n=983$), originating from 210 half and full-sib families of the Finnish national rainbow trout breeding programme (owned by the Finnish Game and Fisheries Research Institute, FGFRI), were kept indoors in freshwater at the Tervo Fisheries Research and Aquaculture Station (eastern Finland). Each family was initially kept in 150 l family tanks. In February 2002, after 8 months in the family tanks, the fish were tagged with PIT transponders (Trovan Ltd., Germany) to provide individual identification, and allocated to eight tanks of 20 m^3.

The fish were subjected to either of two diets from May 2002, but for the purpose of the present paper the data are considered as a single group. The diets were identical fishmeal and fish oil based modern commercial diets with differing protein/lipid ratio, both conditions having four replicate test tanks. Diet 1 (wet weight); protein 40% and lipid 33%. Diet 2; protein 53% and lipid 20% with a common average proximate content of; moisture 4.7%, ash 7.4%, crude fibre 0.8%, nitrogen-free extracts 13.2%, phosphorus 12.8 ($g \times kg^{-1}$) and energy 24 ($kJ \times g^{-1}$). Pigment (Carophyl Pink) was added in equal amounts (50 ppm) to both diets.

The fish were slaughtered in November 2003 at a mean weight and length of 2.6 kg±0.6 (SD) and 51.5 cm±3.8 (SD), respectively. The fish were gutted directly after slaughter. They were then stored on ice for 48 h, after which 3 cm thick cutlets were cut manually directly in front of the dorsal fin using a knife. The fat percentage of the muscle was estimated by mid-infrared transmission spectroscopy (MIT) (Elvingsson and Sjaunja, 1992).

2.2. Image acquisition

The cutlets were placed directly on a flatbed scanner (CanoScan LiDE 80, Canon Inc., Ohta-ku, Japan), front-down with the abdomen cavity facing up in the image and parallel to the *y*-axis. The cutlets were scanned with a spatial resolution of 200 points per inch, equivalent to a pixel size of 0.127×0.127 mm^2, with 256 grey levels

in each of the three layers in the RGB colour model. The images were stored in the Tiff image file format. The scanner was calibrated using the white surface included in the scanner lid, and the same calibration was used throughout the experiment. The scanner platen was wiped clean between each scanning using paper tissue and a fat-solvent washing liquid. A light-tight box was placed on top of the scanner to prevent light from external sources reaching the image sensor in the carriage (sliding optical unit including a linear array CCD recording RGB colour values).

The images were also translated to the IHS (Gonzalez et al., 2003) and the CIE1976$L^*a^*b^*$ (Wyszecki and Stiles, 2000) colour models. The KODAK Q-60R2 (2004:01) (Eastman Kodak Company, Rochester, New York, USA) colour target was used to calibrate the scanner's RGB values to the CIE1931 *XYZ* tristimulus values needed to calculate the CIE1976$L^*a^*b^*$-values (for formulae see Wyszecki and Stiles, 2000). The estimated *XYZ* tristimulus values from this calibration correlated highly with the known CIE1931 *XYZ* tristimulus values of the colour target ($r=0.999$, $r=0.998$ and $r=0.999$, respectively), indicating a close match. The L^*, a^* and b^* values represent three axes in the colour model; dark to light, green to red and blue to yellow, respectively. Hue (colour) and Chroma (colour intensity or saturation) values provide an intuitive description of the sample colour, and are therefore often preferred to the a^* and b^* values when describing the colour of a sample; Chroma $= (a^{*2}+b^{*2})^{0.5}$, Hue $= \tan^{-1}(b^*/a^*)$ for $a^*>0$ and $b^*>0$, Hue $= 180+\tan^{-1}(b^*/a^*)$ if $a^*<0$ otherwise Hue $= 360-\tan^{-1}(b^*/a^*)$.

2.3. Manual image analysis

A manual image analysis was performed in order to evaluate the results of the automated image analysis methods presented below. A sub-sample of 50 randomly selected fish were analysed by a qualified technician using the roipoly function in Matlab v6.5 (The MathWorks Inc, Massachusetts, USA). This function was used to draw around the cutlet, the red muscle and the dorsal fat depot using the mouse, creating a manual segmentation of these regions.

2.4. Automatic image analysis

The following image analysis methods were developed in Matlab (including the Image Processing and Statistics toolboxes) and are fully automatic. The first step in the automatic image analysis process is to separate the cutlet region from the background. The cutlet region is red, while the cutlet skin and the lid of the scanner have a dark greyish colour in the scanned images. All pixels in the RGB image whose R value is greater than both the G and B values are predominantly red and can therefore be identified as representing cutlet; all other pixels are identified as background.

$$g_{\text{cutlet}}(x,y) = \begin{cases} 1 & \text{if } \mathrm{R}(x,y) >> \mathrm{G}(x,y) \wedge \mathrm{R}(x,y) >> \mathrm{B}(x,y) \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

This segmentation rule generally worked well on all the images, but in some cases blood spots and particularly dark patches of muscle were falsely segmented as background. This resulted in small erroneous holes in the segmented cutlet region. These holes were typically located close to the centre of the epaxial part of the cutlet. The proposed image analysis method removes these holes by setting all pixels (x,y) in g_{Cutlet}(see Eq. (1)) not belonging to the largest connected region of 0s to 1. Small irregularities in the segmentation of the cutlet region border are then removed by first extracting the pixels in the border of the region (x_1, y_1), (x_2, y_2),…,(x_k, y_k) and then reconstructing the region from only the $28<<k$ first Fourier coefficients of the original border. This procedure for smoothing the border of a region is more thoroughly explained in Gonzalez et al. (2003). The selected number of 28 Fourier coefficients was judged as sufficient to remove minor concavities and

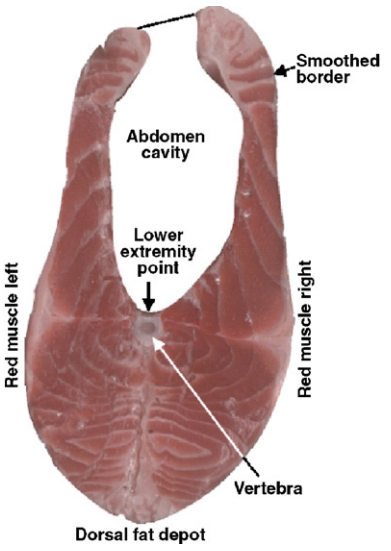

Fig. 1. Result of dominant colour segmentation of one of the scanned cutlets from the image background. The smoothed border is constructed from 28 Fourier coefficients. The positions of the dorsal fat depot, red muscle, vertebra and lower extremity point (LEP) of the abdomen cavity are indicated.

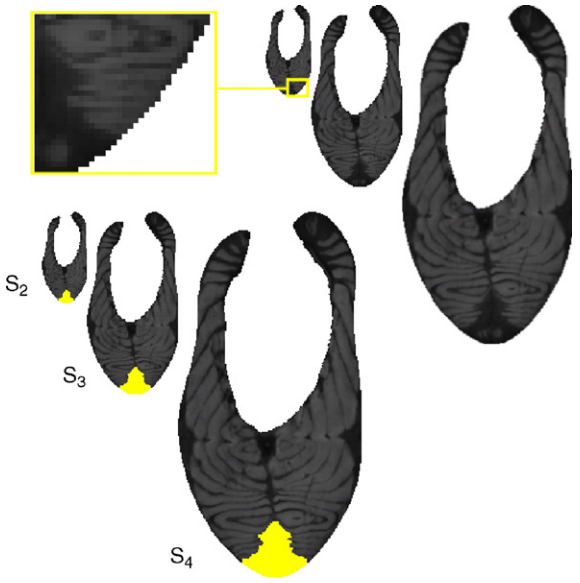

Fig. 2. Pyramidal image structure by wavelet transform of the S colour layer, with or without preliminary segmentation of the dorsal fat depot. Only image levels 1–3 (S_1–S_3) are included in the illustration for simplicity. Notice in the enlarged square from S_1 that the fat stripes are either suppressed or reduced to thin one-pixel-thick structures.

irregularities from the cutlet border while still retaining the overall shape of the cutlet region accurate (Fig. 1). In the following, only coordinates (x,y) where $g_{\text{cutlet}}(x,y)=1$ are included in the analysis.

The dorsal fat depot (DFD), the red muscle (RM), the vertebra (V), and the fat stripes (FS) are all the same white-pink colour in the scanned images (Fig. 1), and therefore cannot be distinguished on the basis of colour values alone. Instead, characteristics such as size, shape and location of these structures have to be utilised by the image analysis methods. At low resolutions massive regions such as the dorsal fat depot and the vertebra remains clearly visible, while the comparatively thin fat stripes are suppressed (Fig. 2). The dorsal fat depot was judged to have especially dark values compared to the cutlet flesh in the S layer of the IHS colour model. The first step towards segmenting the dorsal fat depot is therefore to create a 5-level pyramidal image structure of the S colour layer by a 4-scale wavelet transform (Gonzalez et al., 2003), the first level S_1 having approximately $1/2^4$ the length and breadth of the original image, S_2 $1/2^3$ and so on. The initial segmentation of the dorsal fat depot from the rest of the cutlet is then performed at low resolution.

$$g_{1\ \text{DFD}}(x,y) = \begin{cases} 1 & \text{if } S_1(x,y)<T \\ 0 & \text{otherwise} \end{cases} \qquad (2)$$

It was not possible to find a constant threshold value T in Eq. (2) that would segment the dorsal fat depot correctly in all the images. The threshold therefore had to be set individually for each image. Two methods that automatically find a threshold based on the pixel values in an image are the Otsu thresholding method (Chi et al., 1996) and the Fuzzy thresholding method (Chi et al., 1996; Brandtberg, 2002). Otsu's method finds the threshold T that maximises the between-class variance (function graythresh, Matlab), while Fuzzy thresholding finds the threshold T so that the grey levels of the pixels above and below the threshold are as close to their class average as possible. Of these two methods, Fuzzy thresholding was visually found to give the best segmentation of the DFD region. In some of the images, particularly thick fat stripes are still present at low resolution. The initial segmentation (Eq. (2)) is therefore followed by an opening operation (function imopen, Matlab) with a 3×3-structure element to remove any thin areas of 1's in $g_{2\text{DFD}}$ representing fat stripes. The image analysis finally identifies the DFD region as the largest connected region of 1s in the lower part of the binary image $g_{2\text{DFD}}$. This segmentation forms the basis of a hierarchical thresholding procedure through images of increasing detail (S_1 to $S_5=S$, Fig. 2), until a binary function g_{DFD} representing the dorsal fat depot in the image's original size is obtained. Each step of the hierarchical thresholding procedure consists of dilating (function imdilate, Matlab) with a 3×3 structure element and then iteratively removing pixels from the border of the dorsal fat depot region that do not meet the thresholding criteria (Eq. (2)). The dorsal fat depot region is finally smoothed in the same way as the cutlet region, but

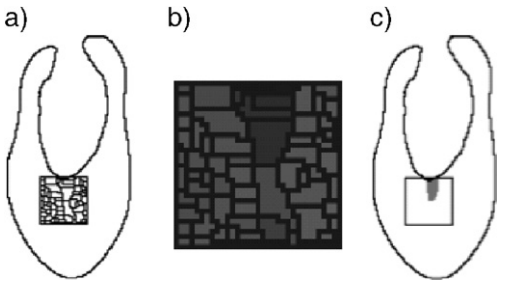

Fig. 3. The image analysis method identifies the vertebra by first defining a small square area of interest (AI) below the lower extremity point of the abdomen cavity (see Fig. 1), performing watershed segmentation on this AI in S_2 (see Fig. 2) and then identifying the watershed region with the darkest mean grey level in S_2. a) The initial watershed segmentation of the AI. b) Average grey level in S_2 for the watershed regions. Notice that the darkest regions represent the vertebra. c) Final segmentation of vertebra at level 2 (indicated as a grey region).

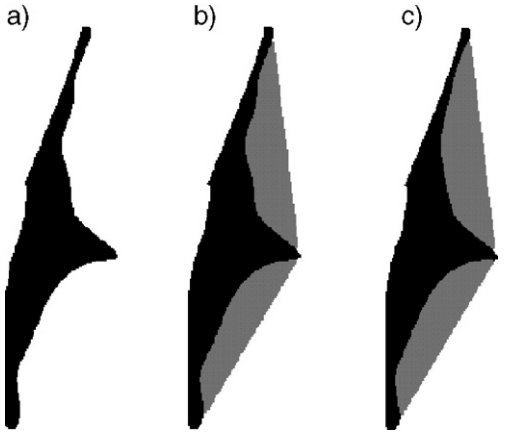

Fig. 4. a) Initial segmentation of a red muscle region. b) Two convex hulls, one above and one below the main promontory. c) Result after removing the convex region of each of the convex hulls in b) from the original segmentation in a).

only using the 14 first Fourier descriptors. A smaller number of coefficients is used, since the dorsal fat depot region has a less complex shape than the cutlet region.

The vertebra is segmented using the same hierarchical segmentation algorithm as described above. However, the initial region cannot be found in the same way. The vertebra is much smaller than the dorsal fat depot. It was therefore necessary to start at a one step larger level in the pyramidal image structure (S_2 instead of S_1). It was also not possible to detach the vertebra region from neighbouring fat stripes by an opening operation as this in some cases also would remove the vertebra region. The initial vertebra region is, instead, identified by first defining a small rectangle below the lower extremity point of the abdomen cavity as the area of interest (AI) (Fig. 3a). Watershed segmentation (function watershed, Matlab) is then applied to this area in S_2, creating a number of smaller regions within the AI (Fig. 3ab). The watershed region with the lowest mean S_2 value is then identified as representing the vertebra (Fig. 3b), including any adjacent regions with similar mean S_2 value (difference $<|10|$) (Fig. 3bc). After the hierarchical segmentation, the vertebra region is finally smoothed by enforcing a border of eight Fourier coefficients.

Similar procedures to those for the dorsal fat depot region and the vertebra region were also tried for the red muscle regions. However, these were visually found to segment the red muscle regions incorrectly in many of the images. Another approach was therefore tried. The red muscle was judged to be especially apparent in the B layer of the RGB colour model. The suggested method consists of first applying a 21×21 median filter to the B layer. This operation removes noise and makes the image regions in the B layer more uniform. A threshold that segments the light red muscle in the smoothed B layer from the dark white muscle is then found by Fuzzy thresholding (see above). The left and right red muscle regions are identified based on their positions; on the right and left extremities of the cutlet region, at the height of the segmented vertebra region (Fig. 5). Preprocessing by means of median filtering removes adjacent thin fat lines, but thicker lines remain as 'bumps' on the initial segmentation of the red muscle region (Fig. 4a). These 'bumps' are removed in three steps; first, the convex hull (the smallest convex region containing the region) (function regionprops, Matlab) of the red muscle region (left or right) is found. The red muscle forms a thin layer close to the edge of the cutlet, and is concentrated into a wedge at the major horizontal septum. This means that two large convex deficiency regions (regions that form part of the convex hull but not the original region) are created above and below the wedge (Fig. 4b). The convex hulls of these two regions are then calculated (Fig. 4c). These two hulls will by design include the

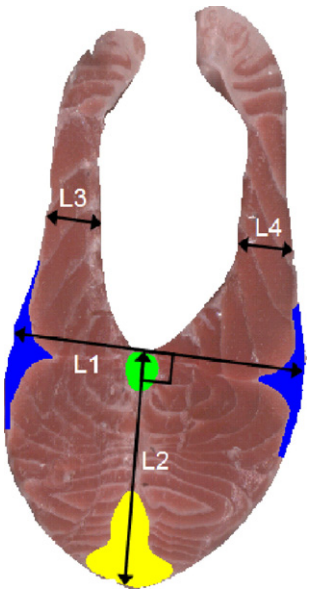

Fig. 5. Final segmentation of the dorsal fat depot (yellow region), red muscle (blue regions) and the vertebra (green region). The line L1 goes through the lower extremity point of the abdomen cavity and is perpendicular to a line L2 going from this point to the upper extremity point of the dorsal fat depot region. The lines L3 and L4 are parallel to L1, and are positioned at half the height of L2 up from the lower extremity point of the abdomen cavity. The Euclidean length of L1 is used as a width measurement, Euclidean length of L2 as the height of the epaxial part of the cutlet, and the Euclidean lengths of L3 and L4 as the width of the left and right belly flaps.

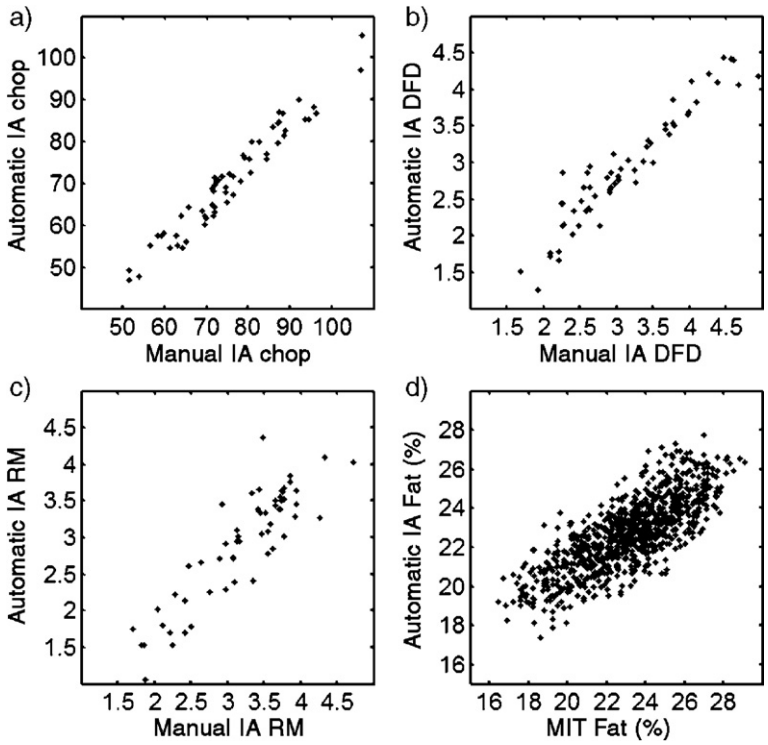

Fig. 6. Results from the automatic image analysis vs. control measurements. A manual segmentation of a) the Area, b) the dorsal fat depot (DFD) and c) the red muscle (RM) was performed on 50 randomly selected cutlets and compared to the results from the automatic image analysis. d) Fat percentage estimated by MIT vs. fat percentage estimated by automatic image analysis. Automatic IA Fat(%)=−27.67+0.47 relAreaFat+0.37 * mean(L^*)+0.50 * mean(Chroma), where relAreaFat is the area of the fat segmented by the *k*-mean method on the B colour layer of the cutlet region and mean(L^*) and mean(Chroma) are the mean values of the L^* and Chroma colour layers, respectively, for the cutlet region.

'bumps'. The last step is therefore to state that all pixels in these two regions are not in the final red muscle region, effectively removing the 'bumps' from the final segmentation (Figs. 4c and 5).

The white muscle region is identified as the remaining region after which the dorsal fat depot, vertebra and red muscle have been removed from the cutlet region (Fig. 5).

$$g_{WM} = g_{cutlet} \cap \neg (g_{DFD} \cup g_V \cup g_{RM}), \quad (3)$$

where g_{WM} is the binary function representing the white muscle region, g_{Cutlet} the cutlet region, g_{DFD} the dorsal fat depot region, g_V the vertebra region and g_{RM} the red muscle regions.

The loins of the cutlet are identified as the two regions above a line going through the lower extremity point of the abdominal cavity (L1, Fig. 5). This line is perpendicular to a line from this point to the upper extremity point of the dorsal fat depot (L2, Fig. 5). L1 divides the epaxial from the hypaxial part of the cutlet.

All the colour layers, except the H and Hue layers, are bimodally distributed, with one of the peaks predominantly representing the grey levels of the muscle and the other the grey levels of the fat. Otsu's method and Fuzzy thresholding were therefore tried on all the computed colour layers to segment the fat stripes from the white muscle. In addition, the *k*-mean method used by Marty-Mahé et al. (2004) was tried. The *k*-mean algorithm starts out with two seed points, m_1 and m_2, defining two classes C_1 and C_2. The grey levels of the colour layer are assigned to the class with the nearest seed. The mean values of the classes are then used as new seed points. This is iterated until the sum of grey level to seed distance is minimised (Khattree and Naik, 2000).

2.5. Output from the image analysis

Output variables from the image analysis methods include three groups of variables; thirteen area variables, eight length variables and one hundred colour variables. The image analysis method calculates the areas of the whole cutlet, the epaxial part (below L1, Fig. 5), the

epaxial part left and right of L2 (Fig. 5), the hypaxial part (above L1, Fig. 5), the loins (left and right), the dorsal fat depot, the total red muscle and the individual red muscle (left and right), the white muscle (including fat stripes) and the area of the fat stripes. The length variables include the height of the epaxial cutlet and the whole cutlet and the width of the cutlet and loins. The height of the epaxial part of the cutlet and the width of the cutlet as a whole are calculated as the Euclidean length of L1 and L2 (Fig. 5), respectively. Then Euclidean length of L1 left and right of L2 are also calculated. The major and minor axes (function regionprops, Matlab) of the cutlet region are calculated as alternative measures of height and width for the whole cutlet. The widths of the loins are defined as the Euclidean lengths of the lines L3 and L4 (Fig. 5), which are parallel to L1 and at a distance (1/2 * length of L2) up from the lower extremity point of the abdominal cavity.

The collected colour values of all the pixels in a region can be described by descriptive statistics. Statistical colour values were calculated for the CIE1976*L*a*b**, Hue and Chroma colour layers for selected morphological regions; whole cutlet, epaxial without dorsal fat depot and red muscle, and epaxial left and right. The statistical descriptors included the mean value, trimmed mean (mean value of remaining observations after all observations below the 5 and above the 95 percentiles have been removed from the original set of observations), 25th, 50th and 75th percentiles.

2.6. Statistics

The statistical procedures were performed using the SAS software package v.8.0.2 (SAS Institute Inc., Cary, North Carolina, USA). Spearman correlations between manual and automatic segmentation of the whole cutlet, the dorsal fat depot and the red muscle areas were calculated (proc corr, SAS) in order to evaluate the performance of the automatic image analysis. The percentage of segmented fat (fat stripes and dorsal fat depot) in the cutlet region was compared to the fat percentage measured by MIT (proc corr, SAS). In order to assess whether the predictive power of the image analysis could be further improved, the areas of the two fat regions were combined with colour variables in multi-trait regression models (proc reg, SAS). Several of the combinations resulted in improved correlations with MIT fat percentage. One of the best performing combinations is given in the text below.

Principal component analysis (PCA) loading plots were used to find a smaller number of representative variables of special interest for future analysis of the data in relation to biological parameters. One PCA loading plot was developed for each of the three groups of variables; (1) area measurements, (2) height and width measurements, and (3) colour measurements. For this purpose, the areas of the cutlet, loins, dorsal fat depot and red muscle regions were normalised to fish size, to prevent fish size alone from influencing the variables. The area measurements were normalised by multiplying by fish length, thus creating the volume of the region, and then dividing by carcass weight to normalise. The length and width variables were normalised by dividing by fish length. Variables that are close together in a PCA loading plot correlate highly. It may therefore be sufficient to only include one of these variables in further biological studies, as they yield the same (or very similar) information. The normalised variables are always referred to as normalised in the following sections.

3. Results and discussion

3.1. Validation of the image analysis methods

The results from the automatic segmentation methods for identifying the cutlet region, the dorsal fat depot region, and the red muscle regions agreed visually with the actual extent of these regions, and the areas given by the automatic and manual image segmentation of the sub-sample of 50 cutlets correlated highly with each other (Fig. 6abc, $r=0.97$, $r=0.95$ and $r=0.91$, respectively). This indicates that the automatic segmentation of these regions was accurate and provided fast and reliable ways of measuring.

The area of the fat relative to cutlet area correlated well with fat percentage measured by MIT. This result was consistent for both the Otsu and the *k*-mean thresholding methods performed on the B colour layer of the RGB colour model ($r=0.69$ and $r=0.68$, respectively). These correlations are of similar magnitude to the correlations obtained for the fat segmentation by Marty-Mahé et al. (2004) ($r=0.72$). This is a promising result, especially considering that the images of the brown trout cutlets in Marty-Mahé et al.'s study were obtained using an elaborate lighting scheme, while the images in the current study were obtained using a regular consumer flatbed scanner. The current study also included significantly more cutlets ($n=987$ vs. $n=48$).

The chief advantage of Marty-Mahé et al.'s study over our work was probably that the colour vision system with diffuse lighting that they employed minimised specular reflection on the surface of the cutlets, while specular reflections were relatively frequent in the scanned images of the present study, especially when the cutlets were unevenly cut. Specular reflections appear as white

bright areas in the images and may therefore have wrongly been segmented as areas representing fat stripes by the automatic image analysis method. It is therefore likely that even higher correlations would have been achieved in the present study if specular reflections had been avoided during the scanning of the cutlets. Another difference is that the loins were still connected in the study by Marty-Mahé et al. (2004), while the loins were separated during evisceration in the current study, which made it difficult to keep them from leaning over during the scanning procedure. This will probably have affected the result of the segmentation of the fat stripes and may have been particularly important since the fat percentage is higher in the loins than in the epaxial part of a salmonid cutlet (Morris, 2001). However, a much higher correlation between the relative area of the fat regions and fat percentage as measured by MIT than those presented here is probably not to be expected. The fat in the muscle is not only stored in the fat stripes but also in the muscle fibres between the stripes (Borderías et al., 1999). Moreover, the fat percentage of the stripes per area can vary. As a result any method that calculates the fat percentage based on the area of the fat stripes will never be able to achieve a perfect correlation to the actual fat percentage of the muscle.

The highest correlation between image analysis variables and fat percentage in Marty-Mahé et al.'s study was in fact obtained by the mean lightness (L^*) values of the cutlet region ($r=0.77$). This can easily be explained if equal pigmentation of the muscle is assumed; thicker fat stripes and more fat marbling in the muscle tissue will then lead to a more white cutlet surface, and thus to higher mean L^* values. In this study, the correlation of mean L^* value with MIT fat percentage was only $r=0.52$, indicating a less uniformly pigmented population. However, it is clear that both fat area and flesh colour are indicative of fat percentage also in the present study, and by combining these variables in a linear regression model: Automatic IA Fat(%)=−27.67+0.47 relAreaFat+0.37 * mean(L^*)+0.50 * mean(Chroma) it was possible to achieve a higher correlation to fat percentage measured by MIT than for the individual variables alone (Fig. 6d, $r=0.78$), where relAreaFat is the area of the fat segmented by the k-mean method on the B colour layer of the cutlet region, and mean(L^*) and mean(Chroma) are the mean values of the L^* and Chroma colour layers, respectively, for the cutlet region.

The basic requirement of the presented methods is that the muscle tissue has a red hue. 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 (MRI), infrared imaging (IR), 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.

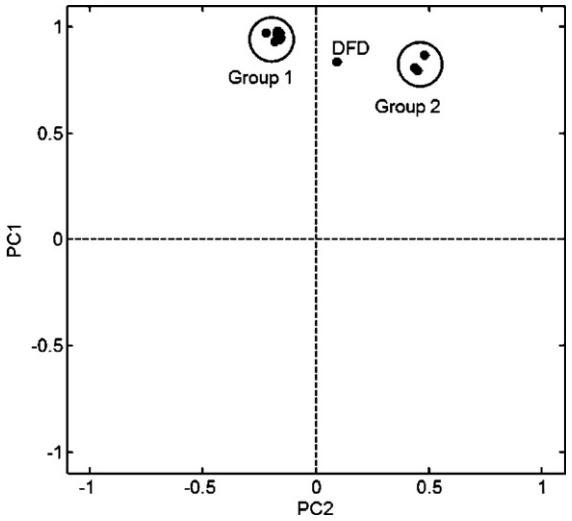

Fig. 7. PCA loading plot for the normalized areas of the segmented regions. PCA 1 explains 85% of total variance, PCA 27%. Group 1: Variables closely linked to normalised cutlet size (n. area of cutlet, n. areas of abdomen walls, n. area of the epaxials part of the cutlet, n. area of white muscle). Group 2: Variables linked to red muscle size (n. area of the red muscle left and right). The n. area of the dorsal fat depot (DFD) is positioned between Group 1 and Group 2.

3.2. Evaluating the image analysis variables

The normalised areas of all the segmented regions all correlated strongly along the first principal component (85% of the variation, Fig. 7). However, there is a division into two main groups along the second principal component (7% of the variation). Group 1 contains variables directly related to the normalised size of the cutlet (normalised area of cutlet, white muscle, abdomen walls left/right, epaxial cutlet left/right). Group 2 includes the normalised areas of the red muscle (both left and right individually and as a whole). The normalised area of the dorsal fat depot is positioned between Group 1 and Group 2. In other words; the main effect on all the area variables came from normalised cutlet area, but the areas of the red muscle and dorsal fat depot were also influenced by other factors. The position of the normalised area of the dorsal fat depot, between Group 1 (cutlet area) and Group 2 (red muscle) can be explained by the fact that large mature fish deposit more fat (Shearer, 2001), and that fish with more red muscle require more external fat. Red muscle has a greater

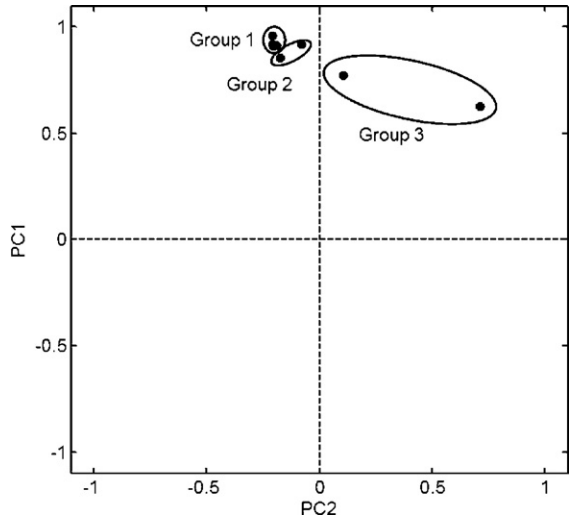

Fig. 8. PCA loading plot of normalized length components, PCA 1 explains 68% of total variance, PCA 2 12%. Group 1: Variables describing the normalised widths of the cutlet regions; n. Euclidean length of L1 and L1 left and right of L2, and the minor axis. Group2: Variables describing the normalised heights of the cutlet regions; n. Euclidean length of L2 and the major axis. Group 3: Variables describing the normalised widths of the belly flaps; n. Euclidean lengths of L3 and L4. The width of the left loin is placed furthest to the left in the group, while the width of the right loin is furthest to the right.

functional need for fat (aerobic metabolism) per gram tissue than white muscle (Kiessling et al., 1991) and a fish with relatively larger red muscle mass might therefore have a greater need for extra muscular fat depots than a fish with a relatively larger white muscle mass. In conclusion, a subset of three normalised area variables from the image analysis (the normalised areas of the cutlet, dorsal fat depot and red muscle) is sufficient to give a description of the main variation.

The normalised lengths of the internal cutlet to the length of the fish are all strongly correlated along the first principal component (68% of the variation, Fig. 8). However, there is a spread along the second principal component (12% of the variation). The normalised widths of the whole cutlet measured as minor axes, |L1|, |L1left|, and |L1right| (see Fig. 5) are clustered together (Group1). The normalised height of the cutlet (major axis) and the normalised height of the epaxial part of the cutlet (|L2|) are also grouped together, but to a smaller degree (Group 2). Group 3 contains the normalised widths of the left and right loins. The epaxial part (Groups 1 and 2) of the cutlet seems to grow symmetrically in height and width. The hypaxial (the normalised widths of the loins) (Group 3), however, is clearly differentiated from the epaxial measurements along PC 2. There is also a clear difference within the hypaxial measurements. The normalised width of the left abdomen wall is positioned near the centre of PC 2, while the normalised width of the right abdomen wall is positioned far to the right. This indicates a certain asymmetry between the widths of the left and right loins. The asymmetry between the right and left loin cannot be explained by that the loins leaned during the scanning procedure (discussed above), as the width of the loins is measured at a position where the ribs are still strong enough to keep the loins in place (Fig. 5). On this basis, the normalised width and height of the cutlet together with normalised widths of the left and right loin are suggested as interesting variables for further biological analysis.

The statistical colour descriptors (mean and percentile values for the L^*, a^*, b^*, Hue and Chroma colour layers calculated for the cutlet region as a whole, the epaxial part of the cutlet, and the epaxial left and right part) are clustered in two main groups along the first principal component (67% of the variation) (Fig. 9); Group 1 contains the a^*, b^* and Chroma descriptors and is highly positively correlated along PC1, while Group 2 contains the L^* and Hue descriptors and is negatively correlated along PC1. This inverse relationship between the two groups is in line with several colorimetric studies of salmonid flesh, which have found a^*, b^* and Chroma

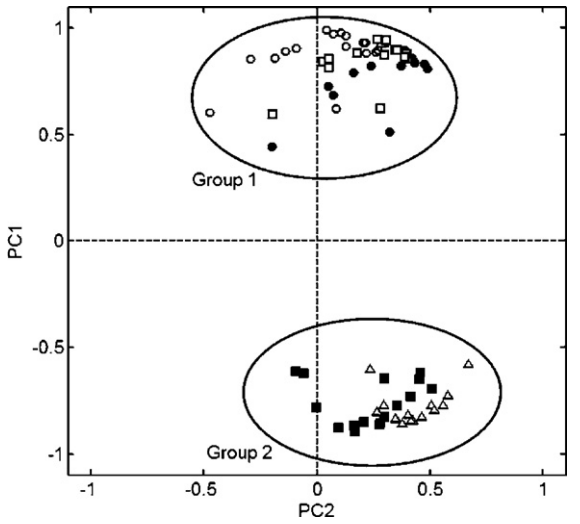

Fig. 9. PCA loading plot for the statistical colour descriptors (mean and percentile values for the L^*, a^*, b^*, Hue and Chroma colour layers calculated for the cutlet region as a whole, the epaxial part of the cutlet (above L1 see Fig. 5), and the left and right of L2 epaxial part. PCA 1: explains 67% of total variance, PCA 2 11%. Group 1 includes the a^* (open circles), b^*(close circles) and Chroma (open squares) descriptors, while Group 2 includes the Hue (closed squares) and L* (open triangles) descriptors.

values to be positively correlated with astaxanthin and the Hue and L^* values to be negatively correlated with astaxanthin (Wathne et al., 1998; Robb, 2001; Baker et al., 2002; Rønsholdt, 2005). For the a^*, b^* and Chroma descriptors in Group 1, the statistical descriptor furthest to the left in the PCA loading plot is the 25th percentile. This statistical descriptor, by its nature, samples a pixel with relatively low colour saturation, and is therefore probably sampling a pixel that represents the white-pink fat (or alternatively muscle tissue with a high fat content). The 75th percentiles are grouped furthest to the right, and the mean, trimmed mean and median in-between. All in all, this suggests that the 75th percentiles are samplings of pure muscle colour, while the mean, trimmed mean and median values describe the flesh colour as a whole (including fat stripes) On this basis, the 75th percentile and the mean statistics for the a^*, b^* and Chroma values are suggested as measures for pure muscle and overall flesh colour respectively. It is, of course, opposite for the L^* values in Group 2. Here the 75th percentiles represent white-pink fat, while the 25th percentile is a measurement of pure flesh colour.

4. Conclusions

Image analysis is a powerful tool for describing the morphological structure of salmonid cutlets. The automated image analysis methods presented here were able to describe quality properties, such as area of the cutlet, dorsal fat depot, red muscle, fat percentage and colour from a large number of scanned images of rainbow trout cutlets. It is also evident that it is possible to produce images of cutlets of adequate quality for image analysis, using a simple flatbed scanner. No elaborate lighting regime is necessary. The image analysis methods presented here therefore have the potential to be used by the research community.

Acknowledgments

This work was carried out with financial support from the Commission of the European Communities, Quality and Life Management of Living Resources Programme, project Q5RS-2001-0994 “Protein and growth efficiency in salmonid selection (PROGRESS)”. It does not reflect the views of the Commission and in no way anticipates the CEC’s future policy on this area. Parts of the method development were also financed by the Research Council of Norway, through project grants nos. 153178/120. We would also like to acknowledge the skilled technical work of Grete Thorsheim and Britt S. Daae at the Institute of Marine Research Matre, Norway and the staff at the Tervo Fisheries and Aquaculture Research Station, Finland, especially Ossi Ritola and Tuija Paananen.

References

Ballerini, L., Högberg, A., Borgefors, G., Bylund, A.-C., Lindgård, A., Lundström, K., Rakotonirainy, O., Soussi, B., 2000. Testing MRI and image analysis techniques for fat quantification in meat science. Nuclear Science Symposium Conference Record, 2000 IEEE, 15–20 October 2000, Lyon, France.

Ballerini, L., Högberg, A., Lundström, K., Borgefors, G., 2001. Colour image analysis technique for measuring of fat in meat: an application for the meat industry. In: Hunt, M.A. (Ed.), Proceedings of SPIE Volume: 4301, Machine Vision Applications in Industrial Inspection IX, April 2001, San Jose’, California, USA.

Baker, R.T.M., Pfeiffer, A.M., Schöner, F.J., Smith-Lemmon, L., 2002. Pigmenting efficacy of astaxanthin and canthaxanthin in freshwater reared Atlantic salmon, *Salmo salar*. Animal Feed Science and Technology 99, 97–106.

Borderías, A.J., Gomez-Guillen, M.C., Hurtado, O., 1999. Use of image analysis to determine fat and connective tissue in salmon muscle. Eur Food Research Technology 209, 104–107.

Brandtberg, T., 2002. Individual tree-based species classification in high spatial resolution aerial images of forest using fuzzy sets. Fuzzy Sets and Systems 132, 371–387.

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

Chi, Z., Yan, H., Pham, T., 1996. Fuzzy algorithms: with applications to image processing and pattern recognition (advances in Fuzzy systems, applications and theory, Vol 10). World Scientific Pub Co. Pte. Ltd, Singapore, pp. 57–61.

Elvingsson, P., Sjaunja, L.O., 1992. Determination of fat, protein and dry matter content of fish by mid-infrared transmission spectroscopy. Aquaculture and Fisheries Management 23, 453–460.

Gonzalez, R. C., Woods, R. E. Eddins, S. L., 2003. Digital Image Processing using MATLAB. Pearson Prentice Hall. Pearson Education Inc. Upper Saddle River, New Jersey, USA.

Kause, A., Tobin, D., Houlihan, D.F., Martin, S.A.M., Mäntysaari, E.A., Ritola, O., Ruohonen, K., 2006. Feed efficiency of rainbow trout can be improved through selection: different genetic potential on alternative diets. Journal of Animal Science 84, 807–817.

Khattree, R., Naik, D., 2000. Multivariate Data Reduction and Discrimination with SAS® Software. SAS Institute Inc, Cary, USA, pp. 393–421.

Kiessling, A., Kiessling, K.-H., Storebakken, T., Åsgård, T., 1991. Changes in the structure and function of the epaxial muscle of rainbow trout (*Oncorhynchus mykiss*) in relation to ration and age. II: activity of key enzymes in the energy metabolism. Aquaculture 93, 357–372.

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

Morris, P.C., 2001. The effects of nutrition on the composition of farmed fish. In: Kestin, S.C, Warris, P.D. (Eds.), Farmed Fish Quality. Fishing News Books, Cornwall, UK, pp. 161–179.

Robb, D.H.F., 2001. Measurement of fish flesh colour. In: Kestin, S.C, Warris, P.D. (Eds.), Farmed Fish Quality. Fishing News Books, Cornwall UK, pp. 298–306.

Rønsholdt, B., 2005. Can carotenoid content in muscle of salmonids be predicted using simple models derived from instrumental colour measurements? Aquaculture Research 36, 519–524.

Rønsholdt, B., Nielsen, H., Færgemand, J., McLean, E., 2000. Evaluation of image analysis as a method for examining carcass composition of rainbow trout. Ribarstvo 58, 3–11.

Shearer, K.D., 2001. The effect of diet composition and feeding regime on the proximate composition of farmed fishes. In: Kestin, S.C., Warris, P.D. (Eds.), Farmed Fish Quality. Fishing News Books, Cornwall, UK, pp. 31–41.

Wathne, E., Bjerkeng, B., Storebakken, T., Vassvik, V., Odlands, A.B., 1998. Pigmentation of Atlantic salmon (*Salmo salar*) fed astaxanthin in all meals or in alternating meals. Aquaculture 157, 297–309.

Wyszecki, G., Stiles, W.S., 2000. Table I (6.5.1) Three-dimensional color spaces and color-difference formulae. In: Wyszecki, G., Stiles, W.S. (Eds.), Color Science, Concepts and Methods, Quantitative Data and Formulae — Second Edition. John Wiley and Sons Inc, New York, USA, pp. 825–830.