A system for ultrasound image segmentation for loin eye measurements in swine

A.W.K. TONG\textsuperscript{1}, R. QURESHI\textsuperscript{2}, X. LI\textsuperscript{2} and A.P. SATHER\textsuperscript{1}

\textsuperscript{1}Lacombe Research Centre, Agriculture and Agri-Food Canada, Lacombe, AB, Canada T4L 1W1; and \textsuperscript{2}Department of Computing Science, University of Alberta, Edmonton, AB, Canada T6G 2H1. Received 7 January 1997; accepted 23 October 1997.

INTRODUCTION

A consideration prior to development of image segmentation methodologies is the evaluation of the need to develop such a technology. Is there a reasonable and useful application for these techniques which can be exploited by either the animal breeder for animal evaluation or the meat processor for carcass evaluation? It has been earlier established (Sather et al. 1996) that if the end point is solely to provide an estimate of lean yield of carcass (i.e., carcass value in context of terms of current grading standards) then current technologies are adequate. However, if the investigator wishes to determine individual components of carcass quality or the distribution of value within a carcass then the current grading standards may be inadequate (Sather et al. 1997). Discussions with both the seed stock producers and with the packer/processor sectors within the swine industry confirm varying degrees of dissatisfaction with the status quo, clearly indicating revisions of the current measures of carcass value are required.

There are various methods of edge extraction and image segmentation that may be useful for detecting the muscle \textit{longissimus thoraces} (LT), commonly known as the loin eye. One such technique for finding edges using masks of different sizes to find a consistent way of labeling the information gathered from these masks has been proposed by Marr (1982). This technique was further refined by using a scale space approach in which Witkin (1984) presented a unified description of the extrema in a signal over a wide range of scales. The essential idea of this approach (Eq. 1) is to represent the original image as a family of derived images \( I(x, y; \sigma^2) \) obtained by convolving the original image \( I(x, y) \) with a Gaussian kernel \( G(x, y; \sigma^2) \) of variance \( \sigma^2 \) (Gonzalez and Woods 1992); where \( \sigma \) is the scale parameter, \( \star \) is the convolving operator and \( (x, y) \) is the coordinate within the image of the pixel. Larger values of \( \sigma^2 \) correspond to images of coarser resolutions. That is, the image is viewed at different resolutions much the same way the human eye will vary its focus on an object to isolate it from its background.

\[
I(x, y; \sigma^2) = I_\star(x, y) \star G(x, y; \sigma^2)
\]

(1)
Recent developments of active contours or 'snakes' (Chow and Murray 1993) can play an important role in finding a contour of interest in the image as shown by Kass et al. (1988). A snake is an energy minimizing spline guided by internal energy functions related to the shape of the curve and external energy functions in the images such as lines and edges. These energy functions result in forces which pull the snakes along the direction of these forces. The goal of a snake is to find an equilibrium state in which these forces are minimized. Segmentation techniques can be combined to make use of both the edge and the region information as suggested by Pavlidis and Liow (1990). Their technique is used to combine the information gathered by edge detection and region growing to obtain a unified method for segmenting the image. Morphological operations have also been quite successful in locating objects from the background and removing 'spikes' and overgrowth from the objects (Haralick and Shapiro 1992).

The general overview of the basic image processing procedures is presented by Foley et al. (1994) and Gonzalez and Woods (1992).

The task objective is to detect the loin eye muscle from the loin in ultrasonic images of pigs for carcass evaluation using an image segmentation technique. The main difficulties are the noisy and low contrast input images from an ultrasound scan of a cross section of a pig.

**MATERIALS and METHODS**

Cross sectional ultrasound images of the longissimus muscle were obtained by an Aloka SSD-210DXXII Echo Camera, with a UST-5021, 3.5 MHz probe head (Aloka Co. Ltd., Tokyo, Japan) at a distance of 70 mm from the mid-line of the back between the third and fourth last ribs. Ultrasound images were digitized using the ImaScan digitizing board and ImaCapt capture program (ImaGraph Corp., 220 Mill Road, Chelmsford, MA).

Capture of the input images was attempted at five different locations along the pig's back. Normal animal movement makes image acquisition difficult and impairs obtaining consistent image quality. Therefore, the input images have low contrast and vary considerably from one image to another in terms of texture and contrast. They also have a high level of noise, which is mostly caused by muscle structure and other tissues. Texture-based segmentation techniques often use a priori knowledge of the texture, which is not possible in our case. Threshold-based segmentation techniques require a relatively stable threshold value from image to image or over a large region within an image. Dynamic images from the live pigs make this criterion difficult to achieve.

The objective is to build a system for detecting the LT as well as developing some general procedures for low contrast image segmentation tasks. The system is based on the principle of modular design (Fig. 1) so that different algorithms can be easily added, revised, and updated (Marr 1982). The design is also based on the principle of graceful degradation (Marr1982), so the program can generate certain results even if the input data are not complete or uniform.

Our image segmentation system is divided into two stages. The first stage generates rough contours of the object of interest (loin eye muscle). The second stage refines these contours using information directly from the original image. The general flow of the scheme is shown in Fig. 1.

**Region growing**

The initial step is intended to generate an approximation of the object of interest. Currently, a predefined centre of an object of interest is supplied to the program. Because of the significant amount of variation in contrast and texture, selecting more than one seed point has proved to be necessary. Therefore, five additional seed points around the initial point but within the object of interest are generated. The region growing process starts from each seed point and these regions are subsequently merged to form a single object. The following subsections describe the two main components of this process: the iterative region growing procedure, and the criteria for joining a new pixel.

![Fig. 1. Ultrasound Image Segmentation - System schema](image)

**Iterative region growing process**

This process is designed in an iterative fashion to provide more control than a recursive process, in the overall growth of a region. The pseudo code for the region growing part is shown in Fig. 2.

The above procedure calls a subroutine CRITERIA
(described later) to determine if a pixel is on the boundary of the region. This routine only examines the neighborhood of the pixel and may not make the same decision as a human observer, who can fully utilize global information and a priori knowledge about the region shape.

For this particular detection project, the LT has a convex oval shape, which allows use of the following efficient scanning mechanism (Foley et al. 1990). A stack, a first-in-last-out list, is used to store the starting points of scan lines. Associated with each starting point, a label is used to indicate whether the pixel above or below it has been processed. If the label says 'below', then only the pixels immediately above the current scan line are added to a stack. Similarly the case for label 'above' will be handled. The time complexity of the iterative region growing process is a linear function of \( n \), the number of pixels being considered for inclusion.

**Criteria for joining a new pixel**

This is the crucial part of the region growing process. Although the LT is not easily distinguishable visually, the LT has a higher coherency as compared to the noise due to the reflection from the bones and other tissues. Therefore, viewing the image at a coarser scale can result in the disappearance of artifacts with lower coherency, while leaving the structures with higher coherency like the LT to appear more distinguishable. Before applying the region growing module, pre-processing (convolving) the input image is pre-processed with a Gaussian low-pass filter. In our experiments, optimal results were obtained when the variance \( \sigma^2 \) of the Gaussian filter is set to 1.5.

The 2-D Guassian filter given by Eq. 2 is constructed in the spatial domain and stored in a two dimensional array of 27 by 27. The discrete representation is adequate and is more efficient than the single dimensional form, as it avoids the computation of the radius from the point \((x, y)\).

\[
h(x, y) = \exp\left(\frac{-x^2 + y^2}{2\sigma^2}\right)
\]

In the merging process, the intensity value of each pixel, pixel \((x, y)\), is compared with a threshold \( T \) presented by:

\[
T = \text{seed} + \text{seedv} + \text{bias}
\]

where:

- \( \text{seed} \) = value of seed pixel,
- \( \text{seedv} \) = variance of a small neighborhood (5 by 5) of the seed pixel, and
- \( \text{bias} \) = parameter that provides additional control.

If \( (x, y) > T \), this pixel is recognized as on the boundary. Initially, \( \text{bias} \) is set to 0. The value of \( \text{bias} \) is then adaptively changed according to the vertical position of the scan line. Because the LT region has an oval shape, the region should be, in general wider in the middle of the image and narrower near the top and the bottom. This knowledge is incorporated in a function \( \text{SPAN} \), which can be easily modified to represent other shapes. If a scan line over-grows or under-grows, compared to the general shape information provided by \( \text{SPAN} \), the CRITERIA routine modifies the value of \( \text{bias} \) by a constant. The value of the constant is determined by the user depending on the noise level in the image. Our initial experiments show that a relatively small constant gives better results in noisy images and can avoid oscillation, but requires more iterations. Once this small constant is set, it should work for most images without manual tuning.

The CRITERIA routine can be easily replaced by similar routines implementing other criterion equations, such as the Centroid-Linkage Region Growing (Haralick and Shapiro 1992) or Hybrid-Linkage Region Growing (Haralick and Shapiro 1992). We have implemented those methods and tested one on synthesized images as well as the actual ultrasound images of the LT. The results indicate that those methods are computationally expensive and not effective enough on real images, although they work quite well on synthesised images.

The time complexity of the region growing phase is a linear function of the number of pixels, \( n \), being considered for inclusion in a region. The pre-processing is currently implemented in the spatial domain, which takes more time than the region growing process. A possible improvement would be an implementation infrequency domain.

The steps of the image segmentation system are illustrated in Fig. 3. Part 3a is the original image; the oval shaped slightly
darker region is the LT. Part 3b is the region growing result, which captured the approximate shape of the LT, but some portions of the LT are missing (‘gaps’ and ‘lakes’), and minor leaking occurs in the top and bottom portions of the image. Correction of these aberrations is attempted by later processes.

**Morphological operation**

Through the combined use of selected morphological operators, extraneous parts of an object, such as sudden spikes, irregularities etc., can be removed after region growing (Haralick and Shapiro 1992) through the use of binary erosion and dilation operations.

Fig. 3. The effect of region growing, morphological operations, curve fitting, and snake process on image segmentation compared to artist’s (technician’s) drawing of animal #1.

Our system employs five cycles of erosion followed by five cycles of dilation, which reduce or entirely remove jags and spikes from the LT contour. In some cases, parts of the region are not fully grown, resulting in small ‘lakes’ inside the LT. These can be corrected later in the curve fitting process. In Part 3b, the top and bottom portions have unwanted overgrowth and become the input to the morphological operation module. Part 3e is the result of morphological operations, which removed the ‘lakes’ and some unwanted ‘leaking’.

**Curve fitting**

The image obtained from the morphological operations was the input to a curve fitting module. The purpose of this process is to select a few points along the outline of the detected region and create an analytical representation of the boundary curve. Our system employs a standard third order Bezier polynomial function (Foley et al. 1994) given by:

\[
x(t) = at^3 + bt^2 + ct + d
\]

Equation 5 represents this function in the matrix form:

\[
x(t) = \Phi M_p G_b
\]

where \( \Phi = \begin{bmatrix} t^3 & t^2 & t & 1 \end{bmatrix} \) and the parameter \( t \) is restricted to \([0,1]\) interval, and

\[
M = \begin{bmatrix}
-1 & 3 & -3 & 1 \\
3 & -6 & 3 & 0 \\
-3 & 3 & 0 & 0 \\
1 & 0 & 0 & 0 \\
\end{bmatrix}
\]

is a basis matrix for the Bezier curve and \( G_b \) is the geometry matrix which specifies the four control points selected by using arc segmentation techniques (Haralick and Shapiro 1992).

Since the contour generated by the Bezier module relies on all of the previous processes, any error introduced in earlier stages can be propagated down to the resulting curve. The ‘snakes’ or ‘active contour’ (described below) is used to further refine the contour shape.

The time complexity of this curve fitting part is a linear function of \( m \) where \( m \) is the number of intervals between 0 and 1 for drawing the curve.

The results of the morphological operations (Fig. 3c) are provided to the curve fitting module (Fig. 3d). While the ‘gap’ at the upper-left portion of the LT (Fig. 3b) remained after processing by the curve fitting module, the small ‘hole’ (Fig. 3b) above the ‘gap’ and some spikes around the boundary were reduced or removed.

**Snakes**

The snake process first selects control points along the object contour detected in the first stage. For each control point, the energy values are computed and the control points are moved in such a way that the energy value is minimized.

The snake relies on two energy functions namely the internal energy and the external energy. The internal energy value is determined by the shape of the snake spline and the external energy is determined by features, such as edges or lines, present in the image. These energy terms exert their forces on the snake and adjust a particular point in the contour so that the forces due to these energy functions would be minimized.

Once the energy models are defined, the goal is to find an optimal solution for minimizing the energy values. Various methods have been compared in achieving this goal including: the Variational Calculus approach, the Dynamic Programming approach, and the Greedy Heuristics Search approach. While the first two approaches would result in a globally optimal solution, they are not suitable in our application due to their high computational and memory cost. This is particularly
critical in the development of the image segmentation system since its application will be 'real-time' with a maximum time frame of less than 10 s. This time frame not only includes the time required for image segmentation, but accommodates the hardware module to capture the images and then report preparation and generation.

Our system uses the third approach since it is capable of handling constraints with much less time and storage requirements (Chow and Murray 1993; Williams and Shah 1992). This approach tries to optimize points one at a time rather than attempting a globally optimal solution. An optimal location of a control point within a small neighborhood is determined by assuming all the other points being fixed. Changes are propagated down the entire length of a snake point by point. An iteration is completed when all points have been optimized. The iteration process will continue until no more than a certain minimum number of points is moved in between successive iterations. As only a single point is allowed to move at any intermediate step and the search space for a single point is much smaller, the optimization can proceed much faster. The time complexity of the Greedy approach is a function of the product of $k_1$ and $k_2$, where $k_1$ is the number of control points along the snake spline and $k_2$ is the size of the neighborhood used.

In case of noisy and low energy images, as in our experiments, the above model for external energy, if directly applied to the original input image, cannot yield suitable results. To remove noise and unnecessary details, a standard lowpass filtering and resolution reduction are introduced (Lindeberg 1994). The external energy of this reduced and smoothed image can be obtained by applying an edge operator. Currently this energy function is not optimal, particularly for the LT detection operation. Research in this area is under progress.

To further explain the methods used in our algorithm, we discuss briefly the result of application on the snake module on the original image (Fig. 3a). The results obtained from the curve fitting module (Fig. 3d) are used as the initial curve to feed into the snake module and the result is shown in Fig. 3e. The curve effectively converges to a more accurate location of the boundary. Figure 3f is an artists interpretation of the desired result obtained by manually tracing the LT from the image Fig. 3a. Because of the low contrast of the noisy images obtained in this study, this interpretation must also be considered an approximation. Contrary to classical criteria for evaluation of a segmentation process, successful elucidation of a segmented object must not be determined solely by 'correct' pixel inclusion into the object, but rather the predictive information of the segmented object towards the desired goal of the image segmentation system. In our context, the question "Does the segmented object provide a superior estimate of carcass value relative to current standards of measurements?" becomes the ultimate criterion for evaluation.

**EXPERIMENTAL RESULTS**

Selected results are shown in Figs. 4 and 5 to illustrate the system and some of its current success and deficiencies associated with poor image quality. In the original images, the back of the pig is on the lefthand side of the image. Thus, these images can be visualized as the cross section of a pig laying on its side with its legs facing to the right.

In general, the images were dynamic in terms of contrast and changes in the shape of LT. Animal movement and posture during image acquisition affected both of these parameters and thus overall image quality.

In most images, the left curve was detected successfully, because it is closer to the pig skin and hence closer to the ultrasound camera (i.e., greater contrast). In Fig. 4, image contrast was low in the lower left quadrant. As a result, the snake module failed to completely detect this boundary. Further, in the top right corner of the image, leakage developed during the region growing process from the LT proper in surrounding tissue. This was only partially attenuated by the following morphological operation and curve fitting modules and as result remained after the snake module.

In cases with extremely poor image contrast (Fig. 5), even the outer edge, just below the subcutaneous fat layers, may not necessarily be successfully determined. The right curve of LT, in general, was more difficult to detect, because ultrasound energy is absorbed as it penetrates through into the loin region.

The image presented in Fig. 5 is of particular interest possibly attributable to muscle movement. While it is not evident in the published image, the original images show contrast degradation expected from motion (i.e. diffuse or
Table I: Prediction of carcass merit using carcass grade probe and live and live animal ultrasound adjusted to constant live weight over 30 live pigs.

<table>
<thead>
<tr>
<th>Regression Model</th>
<th>$R^2$</th>
<th>RSD*</th>
<th>$R^2$</th>
<th>RSD*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hennessey Grade Probe $y_i = \mu + \beta_1 GP + \beta_2 \text{fat} + \epsilon_i$</td>
<td>0.76</td>
<td>13.08</td>
<td>0.79</td>
<td>19.10</td>
</tr>
<tr>
<td>Estimated muscle volume $y_i = \mu + \beta_1 \text{MV} + \beta_2 \text{fat} + \epsilon_i$</td>
<td>0.64</td>
<td>15.89</td>
<td>0.70</td>
<td>22.55</td>
</tr>
<tr>
<td>Estimated muscle volume + fat thickness $y_i = \mu + \beta_1 \text{MV} + \beta_2 \text{FT} + \beta_3 \text{fat} + \epsilon_i$</td>
<td>0.79</td>
<td>12.43</td>
<td>0.89</td>
<td>13.92</td>
</tr>
</tbody>
</table>

* RSD is residual standard deviation

smeared edges and ill defined boundaries). Most interestingly because it is an illustration of an image with poor contrast and contrary to expectation, the inner edges of the image were detected successfully during the growing process. The morphological operations enhanced the outer edges, but the curve fitting operators failed to detect the outer LT boundaries. The snake module was not able to fully compensate for the deficiencies arising from the curve fitting model. However, there was residual information, arising along the subcutaneous backfat layers, in the image that was not fully exploited. The curve fitting and snake modules were able to provide through graceful degradation (Marr 1982) a reasonable approximation of upper edges of the LT. This boundary was not evident to the human eye in the original image.

However, two important issues are illustrated in this image. First, the necessity of acquiring the high quality images. Quality is defined in terms of both a standard posture as well as the application of ultrasound techniques for image acquisition. In other words, the animal must present a consistent posture to assist in providing a uniform conformation of the loin. Second, basic principles of ultrasound image acquisition such as a good coupling of the ultrasound head with the subject are absolutely essential for image segmentation processes to be successful and the skill of the operator in capturing a suitable image. With advances in the overall knowledge of the process, these conditions may be relaxed. Currently proper presentation or posture is required for the capture of suitable images. The expectation may be reasonable with a human subject in a medical context, but it must be recognized that the subjects are animals not necessarily interested in co-operating with the ultrasonographer. Therefore, during the initial development of image segmentation methodologies of ultrasound images, these constraints may be required, but for the practical application of this technology neither of these expectations can be expected to be met.

MODEL VALIDATION

To provide an evaluation of the model, an additional 30 market weight animals (approximately 100 kg) were scanned along the length of the LT. The LT was subdivided into four sections (Fig. 6). The image segmentation procedures were applied to the LT between the third and fourth last ribs (current carcass grade site). The same images were evaluated by a technician. Ultrasound images failed to satisfactorily identify the LT at each terminus due to the complex underlying muscle and bone structure. For the purpose of calculating volume, the terminus at each end of the LT was taken to have zero area. Areas were estimated at the centre of the LT and equidistant between the terminus and the centre of the LT. Volumes were then approximated as the sum of each section (Table I). Backfat thickness was also measured. Using this primitive estimate of LT volume, percentage loin of carcass was estimated. Estimated LT volume alone was able to predict 64% of the variation of commercial loin yield of carcass and 70% of lean yield of loin (Fig. 6). When fat thickness was added as auxiliary information to the model, the $R^2$-value increased to 0.79 and 0.89, respectively. For comparative purposes, the electronic grade probe was able to predict 72% of variation of loin yield of carcass and 77% of lean yield of loin. These preliminary results suggest real-time ultrasound can successfully be used to measure loin area, which in turn can be used to provide a measure of loin content of the animal as well as provide estimates at least equivalent to those measured on the carcass.

Fig. 5. The effect of region growing, morphological operations, curve fitting, and snake process on image segmentation compared to artist's (technician's) drawing of animal #3.

TONG, QURESHI, LI and SATHER
Fig. 6. Schematic view of longissimus thoraces illustrating the measurement sites used to estimate loin volume.

DISCUSSION

The system presented in this paper is a step towards integrating various segmentation techniques and constructing a flow of information between these techniques to provide advanced measures of carcass merit over that presently used for carcass grading. The system is designed for achieving some meaningful segmentation in ultrasound scans of a pig's loin. The procedures employed can be applied to other low contrast images. The use of scale space approach in the region growing and in the snake proves to be quite powerful. Refinement of the snake energy function is currently being pursued. Our experiments provide encouraging results. Although the ultrasonic equipment is inexpensive and safe to use, when compared to CT (Computed Tomography) or MR (Magnetic Resonance) (Kallweit 1991), does not give as high an anatomical clarity. Also the dynamic range of ultrasound images is much lower than that available from the CT or MR. As this is a preliminary investigation, illustrating the concept of real-life image analysis for the live animal evaluation, refinement of the algorithms would be expected to provide further improvement in the results obtained from this study. However, the greatest challenge now appears to reduce the variation in image quality associated with image acquisition.

CONCLUSION

The integration of various segmentation techniques and the construction of a flow of information between these techniques can be used in real ultrasound image analysis for live animal evaluation. These same techniques can have potential application for carcass evaluation. It appears that advanced image analysis techniques have the potential for use by both the animal breeder in animal improvement programs and by the packer/processor for establishing equitable carcass value.

ACKNOWLEDGMENTS

The work was funded in part by the Alberta Agriculture Research Institute, the Alberta Pork Producers Development Corporation, and Alberta Swine Genetics. The authors particularly extend their appreciation to Mr. Sam Harbison whose encouragement during this project was most helpful and for the technical support of Mr. Jeff Colyn and Ms. Sophie Zawadski who made and interpreted the ultrasound images. Mr. Norm Shannon, Lethbridge Research Centre, Agriculture and Agri-Food Canada provided much of the initial technical training for the use of the ultrasound equipment.

REFERENCES


