Fat Estimation in Beef Ultrasound Images Using Texture and Adaptive Logic Networks

James Darrell McCauley* Brian R. Thane† A. Dale Whittaker‡

Abstract

Overviews of Adaptive Logic Networks and co–occurrence image texture are presented. These tools are used for both prediction and classification of intramuscular fat in beef from ultrasonic images of both live beef animals and slaughtered carcasses. Results showed that Adaptive Logic Networks perform better than any fat prediction method for beef ultrasound images to date and are a viable alternative to statistical techniques.
Keywords. Meat, Grading, Automation, Ultrasonic, Images.

Instrument Grading of Beef

In the United States, beef carcasses are subjectively evaluated by skilled human graders. One of the primary factors in determining beef grades, and thus market value, is the amount of intramuscular fat, or marbling. The method used by human graders to determine marbling levels involves visual inspection of a cross–sectional area of the *longissimus dorsi* (ribeye) muscle. This process and these graders, though highly trained and skilled, are prone to errors. According to Cross et al. (1980), grading errors can be as high as 20%. From a study done in the United States, the average national error in assigning USDA quality grades to beef carcasses was 7.3%, while a national panel of highly qualified beef graders had an error of 5.5% (Cross et al., 1984) . This indicates that little improvement can be made under the current system. Since the quality grade is one of the most important factors in determining carcass value, the beef industry needs objective methods to measure the quality of beef carcasses. To meet these needs, a major research effort among animal scientists and engineers has been undertaken in this area. For example, Lenhert and Gilliland (1985) describe the design and operational capabilities of a video image analyzer for beef grading as used by Cross et al. (1983) and Wassenberg et al. (1986); Mitchell et al. (1987) describe the use of NMR imaging of pork tissues; Park (1991) describes the design of an ultrasonic probe specifically for the optimal measurement of beef; and Thompson (1991) applied an X–ray technique using computer–aided tomography (CAT) to quantify intramuscular fat content in beef. Also, Forrest et al. (1989) investigated the possibility of combining the technologies of ultrasonic imaging, electromagnetic scanning, and the use of optical electronic fat–lean probes for on–line carcass composition measurements. As the development of objective beef grading technology progresses, individuals trained in the engineering, computer science, and artificial intelligence disciplines, with a background in meat science, are going to play critical roles in identifying the most promising techniques capable of being implemented into the industry's scheme of beef grading.

Ultrasound

As a result of early studies (Cross et al., 1989), ultrasound has been chosen as the tool for developing these objective methods. One of the reasons for the selection of ultrasound is the characteristics of fatty tissue — it is a very good reflector of ultrasound (Krautkramer and Krautkramer, 1983; Park, 1991) . Ultrasound was also chosen due to its relatively low cost of operation, ease of use (little preparation required), and inherent safety. The application of ultrasound to a test specimen is non–invasive, that is, it does not harm the biological tissue. Due to its non–ionizing emission of sound waves, significant biological effects in mammalian tissues have yet to be confirmed (Kremkau, 1989).

Ultrasonic energy, in the context of physiological analysis, is usually type–classified into various modes. In A–mode ultrasound, the amplitude of the reflected signal is recorded over time. In B–mode, the amplitude is represented by a gray–level in an image. A–mode is analyzed using signal processing while B–mode data is analyzed using image processing techniques.

Automation – A Look at the Future

A fully automated, robotic system for beef grading is conceivable. Such a system must be capable of operating

*USDA Fellow, Dept. of Agricultural Engineering, Purdue University, mccauley@ecn.purdue.edu

†Graduate Student, Dept. of Agricultural Engineering, Texas A&M University, thane@diamond.tamu.edu

‡Assistant Professor, Dept. of Agricultural Engineering, Texas A&M University dale@diamond.tamu.edu

Shorter version submitted to Transactions of ASAE.

under the wet conditions of a slaughter floor, should operate in real time, and should be able to determine grades at chain speed (typical range of the number of head of beef slaughtered per facility per day: 350 to 600 hd/hr = 10 to 6 sec/hd.) The major areas of design, where speed is an issue, would be carcass and probe movement and grade determination. Most current research deals with grade determination. McDonald and Chen (1990) present two methods of image segmentation which separate and isolate the *longissimus dorsi* muscle in digital imagery. The ability to isolate this area of interest in imagery reduces the strict requirement of accurate placement of the ultrasonic probe. Feature extraction and estimation of a marbling score should take as little time as possible to allow time for probe maneuvering, carcass movement, and recording of the grade.

Neural Approaches

With the maturation of VLSI circuit technology, neural networks can be implemented in hardware. Operating in parallel, neurocomputers offer tremendous advantages in computation speed over conventional computers requiring algorithmic programming and would be sufficiently fast to implement a real–time beef grading system. Before implementation in neurocomputers, network architectures suitable for the particular problem domain must be simulated in software.

Artificial neural network approaches using three different paradigms for beef grading using ultrasonic A–mode data have already been investigated (McCauley, 1991; Whittaker et al., 1991) and were shown to have potential for marbling classification. Preliminary investigations (Thane, 1991) using a feed–forward architecture and backpropagation learning show that neural approaches for B–mode data may be feasible. Other architectures and learning models, though, must be investigated for B–mode data.

Adaptive Logic Networks

An Adaptive Logic Network (ALN) is a type of neural network built using binary trees (fig. 1) whose nodes represent logic gates. Inputs of real or integer values are encoded (Smith and Stanford, 1990), producing logical or bit vectors, before actually being presented (along with their complements) to the leaves of the tree (Armstrong and Gecsei, 1979). Real numbers are encoded by using "quantization levels" — the greater the number of levels, the more accurate the encoding. Each tree outputs one bit. To produce non–boolean outputs, several trees are used in parallel and the resulting bit vector is decoded.

Training of ALNs consists of assigning the appropriate boolean function (one of AND, OR, LEFT, or RIGHT) to non–leaf nodes. These are initially assigned at random. A simple, heuristic training (Armstrong et al., 1991) solves the credit assignment problem by determining which nodes are responsible for the tree's output. Two counters in each node, which determine the function of the node and which output is more frequently desired, respond to 1-0 and 0-1 input pairs, respectively, when the node is made heuristically responsible. Armstrong et al. (1991) give a more complete description of this algorithm.

ALNs have certain speed advantages over other networks. In a software implementation (or a sequential hardware implementation), most of the computations can be left out. For example, for an AND function, if one input is zero, there is no need in computing the other input. In a hardware implementation, parallel computations done in a tree of combinational logic gates offer extraordinary speedups over ordinary computers. "Comparisons with a recent chip using standard techniques suggest that hardware based on adaptive logic could evaluate functions at least one thousand times faster, and would be in the trillion connections per second range" (Armstrong, 1991). Hardware implementations have already been described (Armstrong, 1976).

Feature Extraction

To use every pixel of an image as input to a neural network is usually not feasible (e.g., an 80×80 image equates to 6400 input nodes). Therefore, the first step in many image analysis problems is selection of an area of interest and feature extraction. Often, the problem domain delegates which features are useful.

Image Texture

Image texture has long been used in image analysis for segmentation and classification. Several approaches have been taken to estimate texture. One of these is through the use of *co–occurrence matrices.* Haralick et al. (1973) compiled a list of 28 textural features which could be computed from co–occurrence matrices. A subset of these features were used to successfully classify various types of imagery (e.g., photomicrographs of rock and satellite imagery). These co–occurrence matrices were later used by Shearer and Holmes (1990) for plant canopy identification and also by Han and Hayes (1990) to estimate percent soil cover.

Co–Occurrence Matrices

Co–occurrence matrices measure the probability that a pixel of one gray–level will be positioned at particular distance and orientation from a pixel of another gray level and will be represented by $P_\delta(i,j)$ for an $n \times n$

Shorter version submitted to Transactions of ASAE.

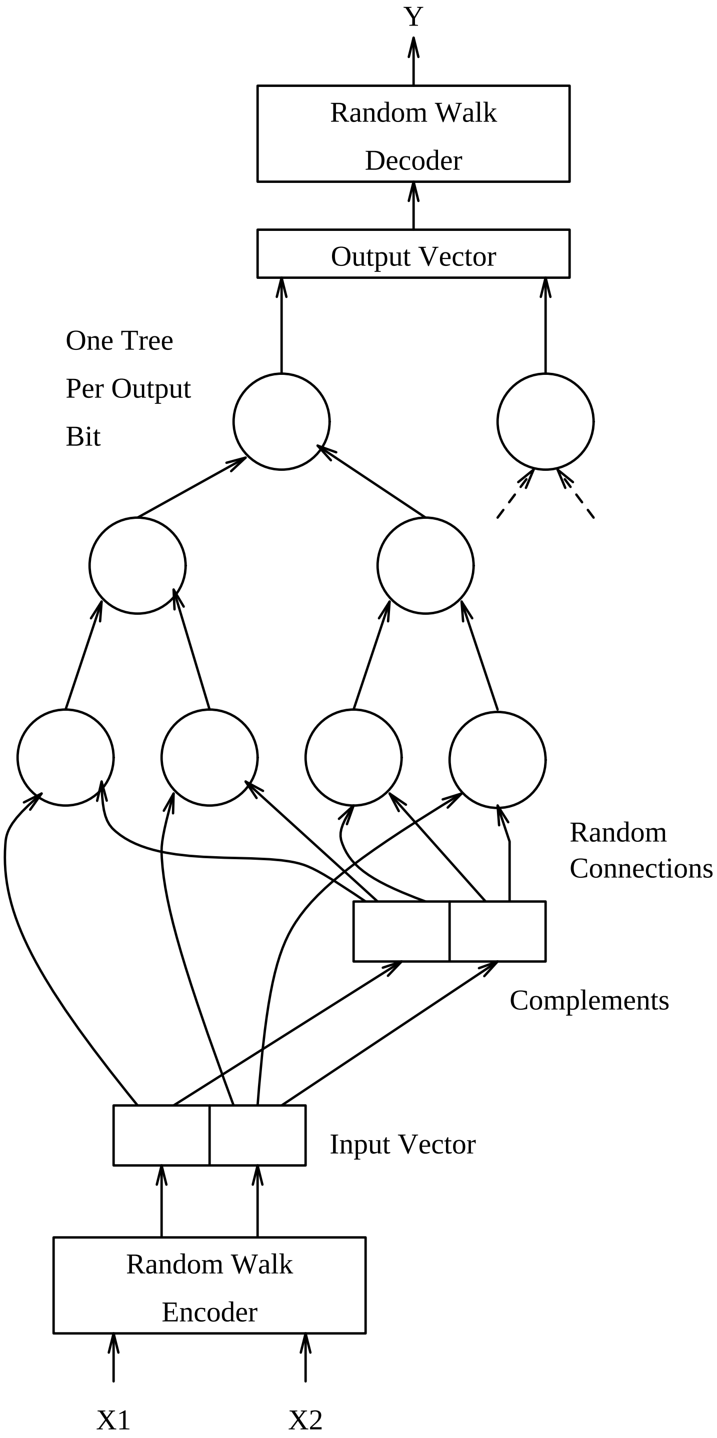


Figure 1 – Structure of adaptive logic networks (Armstrong, 1991).

neighborhood size, where $n \in \{3, 5, 7, \ldots\}$. A directional vector, $\delta = \pm(\delta_x, \delta_y)$, represents the spatial relationship between two pixels. With the first pixel of gray level i at the center of a neighborhood, δ_x represents the horizontal distance (in pixels) to a second pixel of gray level j. Likewise, δ_y represents the distance in the vertical direction. The maximum magnitude of these distances, $\max\{|\delta_x|, |\delta_y|\}$, is always equal to $n - 2$.

A pair collinear rays (δ) from the center of the neighborhood to outer pixel center points (fig. 2), represent the spatial relationship of a co–occurrence matrix for a specified neighborhood size.

A co-occurrence matrix, outlined below,

$$P_\delta = \begin{array}{c} \\ g_1 \\ g_2 \\ \\ g_3 \\ g_n \end{array} \begin{array}{c} \begin{array}{cccc} g_1 & g_2 & g_3 & N_g \end{array} \\ \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1N_g} \\ c_{21} & c_{22} & \cdots & c_{2N_g} \\ \vdots & \vdots & \ddots & \vdots \\ c_{N_g1} & c_{N_g2} & \cdots & c_{N_gN_g} \end{pmatrix} \end{array}$$

would be a symmetric integer matrix for an image with N_g gray levels. The dimensions of this matrix correspond to the number of distinct gray levels in an image. The elements along the main diagonal will always be even integers. For example, element c_{12} would represent the number of co–occurrences of gray levels 1 and 2 for the given spatial relationship δ.

An example image I and the four co–occurrence matrices for an 3×3 neighborhood are shown below.

$$I = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 3 & 3 & 3 \\ 3 & 3 & 3 & 3 \end{pmatrix}$$

$$P_{\delta=\pm(1,0)} = \begin{array}{c} \\ 0 \\ 1 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 3 \end{array} \\ \begin{pmatrix} 4 & 2 & 1 \\ 2 & 4 & 0 \\ 1 & 0 & 10 \end{pmatrix} \end{array} \qquad P_{\delta=\pm(0,1)} = \begin{pmatrix} 6 & 0 & 2 \\ 0 & 4 & 2 \\ 2 & 2 & 6 \end{pmatrix}$$

$$P_{\delta=\pm(1,-1)} = \begin{pmatrix} 2 & 1 & 3 \\ 1 & 2 & 1 \\ 3 & 1 & 4 \end{pmatrix} \qquad P_{\delta=\pm(1,1)} = \begin{pmatrix} 4 & 1 & 0 \\ 1 & 2 & 2 \\ 0 & 2 & 6 \end{pmatrix}$$

Note that these four matrices comprise all possible co-occurrence matrices for $n = 3$. Also note that the co–occurrence matrices are scale invariant — the matrices would be the same if all pixels of gray level 3 in I were changed to 4.

Matrices P_δ are usually normalized to p_δ, according to δ, by the number of nearest neighbor pairs. That is,

$$p_\delta = \frac{P_\delta}{R} \tag{1}$$

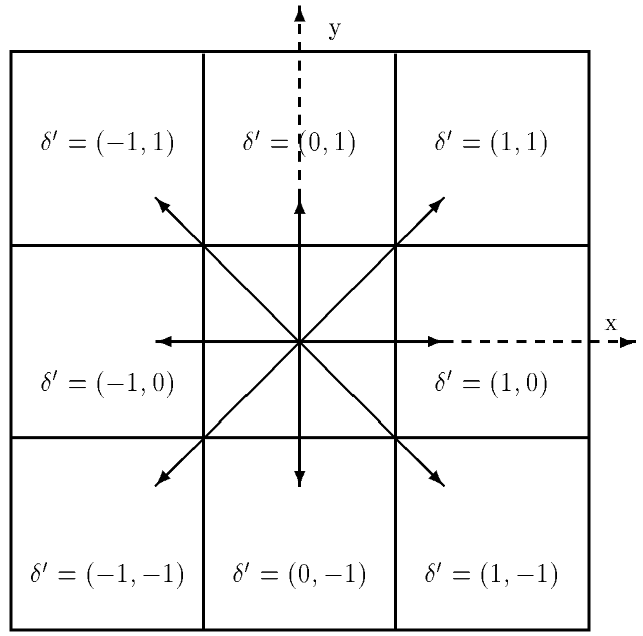


Figure 2 – Position operators for $n = 3$. A directional vector, δ, is comprised of $\pm(\delta_x, \delta_y)$.

where

$$R = \begin{cases} 2N_y(N_x - 1), & \text{if } \delta = \pm(1,0) \\ 2N_x(N_y - 1), & \text{if } \delta = \pm(0,1) \\ 2(N_x - 1)(N_y - 1), & \text{if } \delta = \pm(1,-1) \\ & \text{or } \delta = \pm(1,-1) \end{cases}$$

and

$$N_x, N_y = \text{dimensions of the image.}$$

Given a normalized matrix representing the co–occurrence of gray levels, Haralick et al. (1973) defined 14 texture measures.

The formulae for these measures are given in the next section.

Definition of Textural Features

Haralick et al. (1973) defined 14 measures whose means and ranges comprised 28 features. Each of these statistical measures will be given mathematically; a description or explanation will be given when known. Equations are from Haralick et al. (1973); sources of descriptions will be noted.

Notation

The following notation will be used to define the textural measures:

$p\,(i,j)$ = the (i,j)th entry in a normalized co–occurrence matrix.

$p_x\,(i)$ = the ith entry in the marginal probability matrix obtained by summing the rows of $p\,(i,j)$.

N_g = the number of gray levels in the image.

$$p_y\,(j) = \sum_{i=1}^{N_g} p\,(i,j)$$

$$p_{x+y}\,(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p\,(i,j)\,, \quad i+j = k = 2, 3, \ldots, N_g$$

$$p_{x-y}\,(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p\,(i,j)\,, \quad |i+j| = k = 2, 3, \ldots, 2N_g$$

$$\mu_x = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} ip\,(i,j)$$

$$\mu_y = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} jp\,(i,j)$$

$$\sigma_x = \sqrt{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} i^2 p\,(i,j) - \mu_x^2}$$

$$\sigma_y = \sqrt{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} j^2 p\,(i,j) - \mu_y^2}$$

Textural Features

Angular Second Moment is a measure of homogeneity of the image (Haralick et al., 1973) and is given in the following equation:

$$f_1 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} [p\,(i,j)]^2\,. \tag{2}$$

Higher values indicate fewer amplitude or intensity changes in the image. Fewer changes result in a much sparser co–occurrence matrix.

Contrast measures local variation in the image.

$$f_2 = \sum_{n=0}^{N_g-1} n^2 \left[\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p\,(i,j) \right],\ |i-j| = n. \tag{3}$$

High contrast values indicate high amounts of local variation.

Correlation is a measure of linear dependencies of intensity values of an image.

$$f_3 = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (ij)\,p\,(i,j) - \mu_x\mu_y}{\sigma_x\sigma_y} \tag{4}$$

For an image with large areas of similar intensities, correlation will be much higher than for an image with noisier, uncorrelated intensities (Shearer and Holmes, 1990).

Variance gives an indication variation of values of image intensity and is calculated by:

$$f_4 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - \mu_x)^2 \, p(i,j) \tag{5}$$

For an image whose pixels all have the same intensity, the variance would be zero. This measure of variance is similar to that obtained from histogram analysis (Shearer and Holmes, 1990).

Inverse Difference Moment is another measure of image contrast.

$$f_5 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{1}{1 + (i-j)^2} p(i,j) \tag{6}$$

Sum Average is calculated by the following:

$$f_6 = \sum_{i=2}^{2N_g} i p_{x+y}(i) \tag{7}$$

Sum Variance is calculated by the following:

$$f_7 = \sum_{i=2}^{2N_g} (i - f_8)^2 \, p_{x+y}(i) \tag{8}$$

Sum Entropy is calculated by the following:

$$f_8 = -\sum_{i=2}^{2N_g} p_{x+y}(i) \log\left(p_{x+y}(i)\right) \tag{9}$$

Sum entropy is a measure of randomness within an image.

Entropy is calculated by the following:

$$f_9 = -\sum_{i=2}^{2N_g} p(i,j) \log\left(p(i,j)\right) \tag{10}$$

Entropy is an indication of the complexity within an image. The more complex the image, the higher the entropy value.

Difference Variance is calculated by the following:

$$f_{10} = \text{variance of } p_{x-y} \tag{11}$$

Difference Entropy is calculated by the following:

$$f_{11} = -\sum_{i=0}^{N_g - 1} p_{x-y}(i) \log\left(p_{x-y}(i)\right) \tag{12}$$

Difference entropy, like the measure of sum entropy and entropy, is also an indication of the amount of orderliness (randomness) in an image.

Information Measures of Correlation are calculated by:

$$f_{12} = \frac{HXY - HXY1}{\max[HX, HY]} \tag{13}$$

and

$$f_{13} = \left(1 - \exp\left[-2.0\left(HXY2 - HXY\right)\right]\right)^{\frac{1}{2}} \tag{14}$$

where

$$HXY = -\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) \log\left(p(i,j)\right),$$

HX and HY are entropies of p_x and p_y, and

$$HXY1 = -\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) \log\left(p_x(i)\, p_y(i)\right)$$

$$HXY2 = -\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p_x(i)\, p_y(i) \log\left(p_x(i)\, p_y(i)\right)$$

These two measures of correlation are described in some detail by Linfoot (1957).

Maximal Correlation Coefficient is calculated according to

$$f_{14} = (\text{Second largest eigenvalue of } Q)^{\frac{1}{2}} \tag{15}$$

where

$$Q(i,j) = \sum_{k=0}^{N_g} \frac{p(i,k)\, p(j,k)}{p_x(i)\, p_y(k)}$$

Refer to Bell (1962) for a discussion of the maximal correlation coefficient.

Objective

The objective of this study was to examine an adaptive logic approach to prediction of percentages of intramuscular fat (marbling) from texture features of ultrasonic B–scan images of beef.

For this study, second–order statistical measures of image texture, extracted from ultrasonic B–scan images of beef animals, were investigated using ALNs for the purpose of predicting percentages of intramuscular fat (marbling). Prior studies in which ultrasound B–mode techniques were used to estimate marbling in beef animals have not shown much success. These studies are discussed briefly by Thane (1992). In some studies, it has been assumed that marbling deposits are resolved in ultrasound images, as evidenced by the use of thresholding routines involving measurement of image intensity (Berlow et al., 1989). Thane (1992) has verified that there is a significant contribution of information contained in the high frequency components of an image. This is a result of backscatter, reverberation, Raleigh scattering, etc. This, along with some *ad hoc* measurements of marbling has led the authors to believe that a significant number of the marbling deposits are not resolved. As a result, texture analysis was performed to measure variations in relative intensity which is represented by co–occurrence matrices as described by Haralick et al. (1973). Layer et al. (1990) suggested that tissue characterization by texture measures may possibly provide a new means by which ultrasound B–scan images can be objectively described. Texture analysis is used to measure patterns formed in ultrasound B–scan images that occur, in large, as a result of backscattered ultrasound signals (unresolved marbling deposits). The emphasis of this study was therefore to apply statistical measures of image texture to train ALNs for marbling prediction.

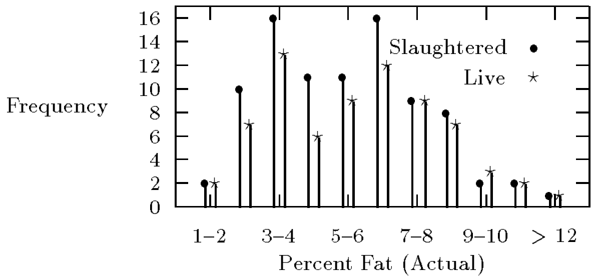


Figure 3 – Frequency distribution of fat for live ($n = 71$) and slaughtered ($n = 88$) animals.

Procedure

Ultrasonic B–scan images from 71 live beef animals and 88 slaughtered beef carcasses (hide–on, pre–rigor) were acquired during the summer of 1990 on the campus of Texas A&M University. The beef specimen consisted of Angus × Hereford heifers and Angus × Simmental steers provided from research stations operated by the Texas A&M University System. See Thane (1992) for a complete description of sampling procedures.

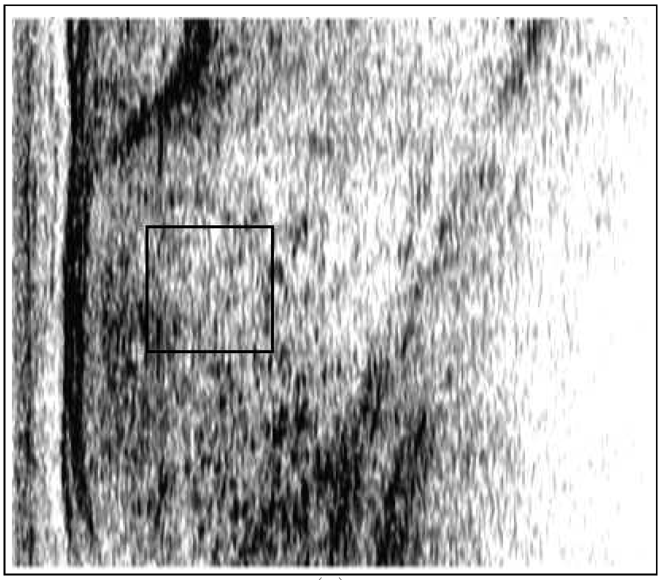
An Aloka Model UST–5024N, 3.5 MHz, linear–stepped array transducer was used, complements of Corometrics Medical Systems, Inc. (Wallingford, CT). During image acquisition, 3 and 5 real-time ultrasonic B–scan images of the *longissimus dorsi* muscle were captured and digitized directly to a computer disk at longitudinal and cross–sectional orientations, respectively. During the scanning process, the 3.5 MHz ultrasonic probe was positioned between and spanning the 12th and 13th rib section of each animal in each instance. Both a Trinitron video monitor (Sony Corp., Park Ridge, NJ) and a Targa–16 image processing board (TrueVision, Inc., Indianapolis, IN) were used during this procedure.

Within 48 hr post–mortem, chemical fat tests (ether extraction) were performed on excised ribeye muscle samples from the same animals to determine percentages of intramuscular fat (AOAC, 1984). Percentages of ether extractable fat were used as endpoint measures to train and validate the capability of the method of marbling prediction (ALN) investigated in this study, since ether extraction is a quantitative measure within a volume. The frequency distribution of intramuscular fat for the beef animals scanned in this study are shown in figure 3.

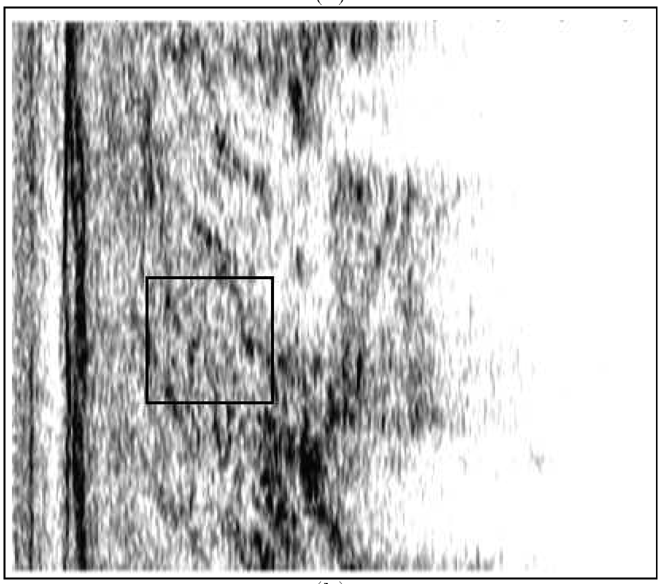
Since each image contained not only portions of the *longissimus dorsi* muscle but backfat, hide, and other anatomical parts not directly related to the desired measurement, an area of interest was selected from the resulting images. Selection of an optimal area of interest was done in another study (Thane, 1992) using statistical analysis of correlation between chemical fat test results and image intensity values of four cropped regions. Using these results, while also considering computational requirements of the different sized regions, the area chosen was the 80×80 region shown in figure 4.

Animals were randomly divided into two approximately equal sets for training and validation (*i.e.*, images for a particular animal would not appear in both training and validation sets). For training sets, all images available for the animal were used. For the validation sets, only one image from the original 3 or 5 images were used so that the neural network made only one prediction per animal. The first thirteen texture features (Eqns. 2–14) were calculated for the area of interest shown in figure 4. The fourteenth feature (Eqn. 15) was excluded since the method for calculating the eigen values would sometimes fail to converge. Feature extraction was done for both cross–sectional and longitudinal probe orientations of live and slaughtered animals, giving a total of four pools from which to select training and validation sets for neural network analysis. The random assignment of animals to training and validation set selection was replicated five times for each of these four data sets.

Software used for ALNs was `atree` release 2.0, Adaptive Logic Network (ALN) simulation program (© 1991. A. Dwelly, R. Manderscheid, W.W. Armstrong). Soft-



(a)



(b)

Figure 4 – Cross–sectional (a) and longitudinal (b) beef images with the 80×80 pixel area–of–interest outlined. The 12th and 13th ribs can be plainly seen in the right portion of the longitudinal image.

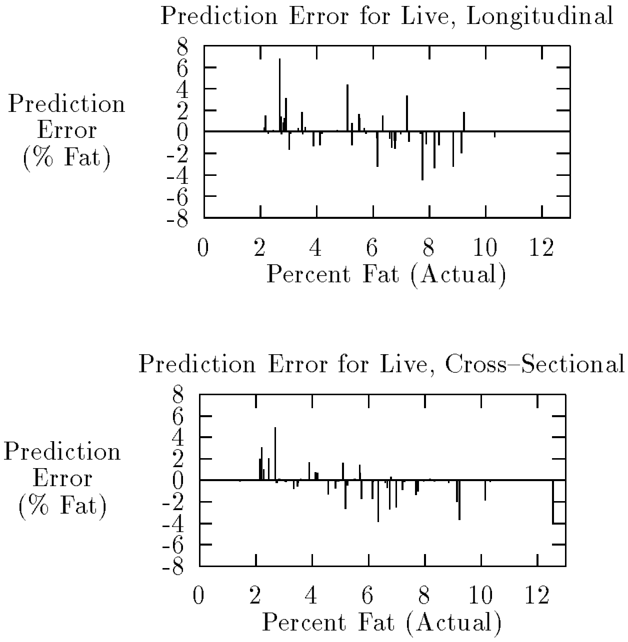


Figure 5 – Distribution of mean (difference) error for live animals.

ware used to calculate texture features was `pgmtexture` (© 1991, Texas Agricultural Experiment Station, employer for hire of James Darrell McCauley), part of the PBMPLUS suite of image tools (© 1991, Jef Poskanzer). Analysis was done using a Sun™ SPARCstation 2.

Results

Prediction and classification results are given in Table 1. Various encoding walks and quantization levels were examined. For both prediction and classification, the best results were for cross–sectional images of live animals. Plots of the prediction error distribution are given in figures 5–6.

For classification (Table 1b), prediction results were divided into three bins of economic importance, corresponding to USDA grade divisions (USDA, 1980). More detailed results of classification are shown in Table 2, showing the distributions of validation sets. Note that more samples lie the middle class of 3–7% fat (see also fig. 3).

Discussion

"Quantization" in Table 1 refers to the number of bins used to divide the real number line for inputs and output.

Table 1: Fat classification and prediction accuracies.

Data Set	Train. Pairs	Valid. Pairs	Quantization Input	Quantization Output	Mean Error[a]
Slaughtered[b]	115	86	30	30	0.94%
Slaughtered[c]	218	87	35	30	0.93%
Live[b]	109	68	35	30	0.83%
Live[c]	154	69	30	30	0.96%

(a) Prediction

Data Set	Train. Pairs	Valid. Pairs	Quantization Input	Quantization Output	Classification Accuracy[d]
Slaughtered[b]	115	86	30	30	73.3%
Slaughtered[c]	192	87	20	30	77.0%
Live[b]	105	66	15	30	78.8%
Live[c]	154	69	15	30	76.8%

(b) Classification

[a]mean error = $\frac{1}{N}\sum|\text{actual} - \text{predicted}|$ percent fat.
[b]Cross–sectional probe orientation.
[c]Longitudinal probe orientation.
[d]Bins corresponding to < 3, 3–7, and > 7% fat.

Table 2: Classification results and distribution of validation data.

Actual Classes	Predicted Classes $< 3\%$	3-7%	$> 7\%$	Near[a] Misses	Type II Accuracy
$< 3\%$	13	0	8	0	61.9%
3-7%	6	39	5	6	78.0%
$> 7\%$	0	4	11	2	73.3%
Type I Accuracy	68.4%	90.7%	45.8%	–	73.3%

(a) Slaughtered, cross–sectional probe orientation.

Actual Classes	Predicted Classes $< 3\%$	3-7%	$> 7\%$	Near[a] Misses	Type II Accuracy
$< 3\%$	12	2	7	2	57.1%
3-7%	2	44	4	3	88.0%
$> 7\%$	0	5	11	1	68.8%
Type I Accuracy	85.7%	86.3%	50.0%	–	77.0%

(b) Slaughtered, longitudinal probe orientation.

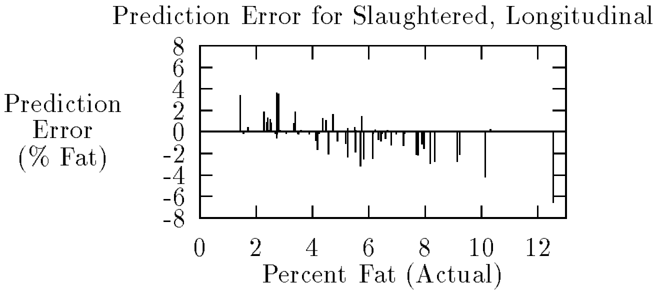
Actual Classes	Predicted Classes $< 3\%$	3-7%	$> 7\%$	Near[a] Misses	Type II Accuracy
$< 3\%$	11	1	2	1	78.6%
3-7%	5	31	1	6	83.8%
$> 7\%$	3	2	10	0	66.7%
Type I Accuracy	57.9%	91.2%	76.9%	–	78.8%

(c) Live, cross–sectional probe orientation.

Actual Classes	Predicted Classes $< 3\%$	3-7%	$> 7\%$	Near[a] Misses	Type II Accuracy
$< 3\%$	12	0	3	0	80.0%
3-7%	7	29	2	6	76.3%
$> 7\%$	1	3	12	2	75.0%
Type I Accuracy	60.0%	90.6%	70.6%	–	76.8%

(d) Live, longitudinal probe orientation.

[a]*Near misses* are defined as a mis–classification by less than $\pm 0.5\%$ fat.



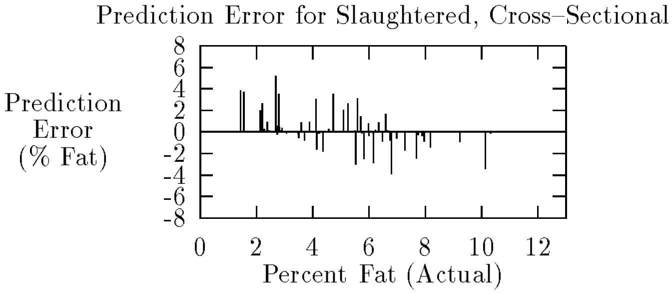


Figure 6 – Distribution of mean (difference) error for slaughtered animals.

For example, the range of fat for live animals was 1.43–12.55% fat (after removing outliers). After dividing this range into 30 intervals (quantization levels), each interval represented 0.37% fat. In this instance, the fifteenth quantization level represented 5.19–5.56% fat. Any sample within this range was considered identical by the network. If the network predicted bin 15, this was decoded to 5.38% fat (the midpoint of [5.19,5.56]). Thus, the best prediction results that could be expected with 30 quantization levels was to predict within 0.37/2 or 0.19% fat. Other quantization levels were examined for input and output variables, but those listed in Table 1 gave the best results. An understanding of the effects of quantization are important when examining the accuracy of ALNs.

In comparison, statistical models presented in earlier studies (Thane, 1992) for this same data predicted with a mean error of 1.38 and 1.16% fat for cross-sectional and longitudinal images, respectively, of slaughtered animals. For images of live animals, mean error in fat prediction was reported to be 1.37 and 1.09% fat for cross-sectional and longitudinal images, respectively. These models not only used image texture features, but incorporated pixel intensity and other measurements. ALNs performed much better than statistical regression on raw images in every instance (Table 1), which was probably because of non–linearities not accounted for in regression studies.

Table 2 shows the distributions of validation sets as well as a more detailed report of classification accuracies. More samples lie the middle class of 3–7% fat (see also fig. 3). This was also the case with training data. Hence, the network was better able to classify samples in this range. Also evident from Table 2 is the network's difficulty with samples of $< 3\%$ fat; if an incorrect classification was made, most often it was predicted to fall in the highest fat class instead of the middle class.

In comparison with other previous studies (Whittaker et al., 1991) classification with ALNs using texture information of image (B–mode) data outperformed neural methods utilizing spectral features of signal (A–mode) data and the same classification scheme.

The general trends shown in figures 5–6, over–predicting for low percentages of fat and under–predicting for high percentages of fat, were consistent with those reported by Thane (1992). The exact reason for this general trend is unknown. Recall that the majority of animals had a fat content near the middle of the range of values (fig. 3). This distribution may be the reason for samples with lower and higher fat content to be not as accurate, in which case a larger data set would improve the networks' predictive ability.

The trend of over–predicting for low percentages of fat and under–predicting for high percentages of fat may also be the result of image resolution and choice of neighboorhood size when building co–occurrence matrices. If the spatial fat distribution in lower marbling level samples consisted of very small blobs and strings of fat, then the resolution of the original image may not have been adequate. Also, recall that Thane (1992) verified that much of the fat is not resolved. Assuming the spatial fat distribution in higher marbling level samples consists of larger blobs and strings of fat, the choice of neighboorhood size used when creating co–occurrence matrices (max $\{|\delta_x|, |\delta_y|\}$) may have not been optimal for samples with higher marbling levels. In other words, samples with low and high fat contents may be more distinguishable if the neighboorhood size was increased. This suggests that using a feature vector with measures calculated for different co–occurrence neighboorhood sizes may give better results.

Regarding conjectures of the size and spatial distribution of fat globs and strings, it should be noted that little work has been done to quantitatitively characterize these properties. McDonald and Chen (1990) have studied geometric modeling of marbling in surface views of beef, but an assumption of linear relationships was made for shape, size, and spatial distribution of fat globules in their work. Though this assumption seems reason-able, this is clearly an area for future work in which co–occurrence image texture could be used.

Incorporation of other image features in the input vector may improve the prediction ability of ALNs. Temporal compounding and autocorrelation with multiple frames and high pass image filtering before feature extraction may also improve the models' predictions (Thane, 1992). Aggregate feature vectors from different probe orientations could potentially increase performance, but the added time requirement to a real–time quality grading system may make this impractical. Aggregate vectors from multiple areas of interest, though, are another possibility. The resolution and consistency of the ultrasonic image is a critical factor for any prediction scheme, but with the data acquisition environment and limitations of current technology, improving this is difficult. Additional data is also expected to improve any approach.

Overall, these results are very encouraging. ALNs have outperformed all prediction methods to date and have proven to be a viable alternative technique to traditional statistical methods. Their potential for extremely fast and adequately accurate performance in a hardware implementation make ALNs attractive. Before implementation, though, a much larger set of data is necessary to train a network that is safe for commercial use.

Conclusions

Adaptive Logic Networks performed much better at predicting intramuscular fat (marbling) in ultrasound B–mode images of beef cattle using image texture information than multivariate statistical regression. By finding additional features to incorporate into the input vector and increasing the quality of images, better performance is expected.

Acknowledgments

The authors thank Dr. Bo Soon Park of Texas A&M University for sharing technical knowledge of ultrasonic measurement and testing, the Texas Agricultural Experiment Station for funding, the Department of Animal Science at Texas A&M University and Dr. Dan Hale of the Texas Agricultural Extension Service for various contributions to the study and expert knowledge in beef grading, and David Smith, an undergraduate in the Department of Agricultural Engineering at Texas A&M University, who aided in collection of ultrasound data, which is the property of the Texas A&M University System and the State of Texas. The authors also acknowledge Academic Computing Services of Texas A&M University and the Engineering Computer Network at Purdue University for use of their computer equipment and Corometrics Medical Systems, Inc. for use of their ultrasonic imaging unit.

References

AOAC. 1984. *AOAC Handbook,* 14th ed. Washington, DC: Association of Official Analytical Chemists.

Armstrong, W. W., A. Dwelly, J. D. Liang, D. Lin, and S. Reynolds. 1991. Learning and generalization in adaptive logic networks, in artificial neural networks. In *Proceedings of the 1991 International Conference on Artificial Neural Networks,* ed. T. Kohonen, K. Makisara, O. Simula, and J. Kangas, volume 2 1173–1176 New York, NY: Elsevier Science Publishing Co., Inc.

Armstrong, W. W. and J. Gecsei. 1979. Adaptation algorithms for binary tree networks. *IEEE Transactions on Systems, Man and Cybertinetics* SMC-9:276–285.

Armstrong, W. W. 1976. Adaptive boolean logic element, U.S. patent 3,934,231. Assigned to Dendronic Decisions, Ltd.

Armstrong, W. W. 1991. *ATREE Documentation,* 2.0 ed. Department of Computer Science, University of Alberta Edmonton, Alberta, Canada.

Bell, C. B. 1962. Mutual information and maximal correlation measures of dependence. *Annals of Mathematical Statistics* 43:587–595.

Berlow, S. M., D. J. Aneshansley, J. A. Throop, and J. R. Stouffer, 1989. Computer analysis of ultrasonic images for grading beef. ASAE Paper No. 89–3569. St. Joseph, MI.

Cross, H. R., L. W. Douglass, E. D. Linderman, C. E. Murphey, J. W. Savell, G. C. Smith, and D. M. Stiffler. 1980. An evaluation of the accuracy and uniformity of the USDA beef quality and yield grading system. Report to the Office of Inspector General, USDA.

Cross, H. R., D. A. Gilliland, P. R. Durland, and S. Seiderman, 1983. Beef carcass evaluation by use of a video image analysis system. *Journal of Animal Science* 57(4):908–917.

Cross, H. R., G. C. Smith, C. Murphey, D. Stiffler, J. W. Savell, and L. W. Douglas. 1984. USDA beef grades: An evaluation of the accuracy and uniformity of their application. *J. Food Qual.* 7:107–120.

Cross, H. R., A. D. Whittaker, and J. W. Savell. 1989. The objective measurement of value in meat animals. In *Proceedings of the Automated Measurement of Beef.* Sydney, Australia: The Australian Meat and Livestock Cooperation.

Forrest, J. C., C. H. Kuei, M. W. Orcutt, A. P. Schinckel, J. R. Stouffer, and M. D. Judge. 1989. A review of potential new methods of on–line pork carcass evaluation. *Journal of Animal Science* 67:2164–2170.

Han, Y. J. and J. C. Hayes. 1990. Soil cover determination by image analysis of textural information. *Transactions of the ASAE* 33(2):681–686.

Haralick, R. M., K. Shanmugam, and I. Dinstein. 1973. Textural features for image classification. *IEEE Trans. on Systems, Man, and Cybertinetics* SMC-3(6):610–620.

Krautkramer, J. and H. Krautkramer. 1983. *Ultrasonic Testing of Materials*, 3rd ed. Berlin: Springer-Verlag.

Kremkau, F. W. 1989. *Diagnostic Ultrasound: Principles, Instruments, and Exercises*, 3rd ed. Philadelphia, PA: W. B. Saunders Company.

Layer, G., I. Zuna, A. Lorenz, H. Zerban, U. Haberkorn, P. Bannasch, G. van Kaick, and U. Rath. 1990. Computerized ultrasound b–scan texture analysis of experimental fatty liver disease: Influence of total lipid content and fat deposit distribution. *Ultrasonic Imaging* 12:171–188.

Lenhert, D. H. and D. A. Gilliland. 1985. The design and testing of an automated beef grader. ASAE Paper No. 85–3035. St. Joseph, MI.

Linfoot, E. H. 1957. An information measure of correlation. *Information and Control* 1:85–89.

McCauley, J. D. 1991. Beef marbling estimation using frequency analysis of ultrasonic a–mode and neural networks. In *Simulation Multiconference Proceedings*, ed. R. J. Uttamsingh and A. M. Wildberger, 23(4): 93–97 San Diego, CA: Society for Computer Simulation.

McDonald, T. P. and Y. R. Chen. 1990. Visual characterization of marbling in beef ribeyes and its relationship to taste parameters. ASAE Paper No. 90–3529. St. Joseph, MI.

McDonald, T. P. and Y. R. Chen. 1990. Separating connected muscle tissue in images of beef carcass ribeyes. *Transactions of the ASAE* 33(6):2059–2065.

Mitchell, A. D., P. C. Wang, and T. H. Elsasser. 1987. Nuclear magnetic resonance imaging of the pig and spectroscopy of pork tissue. *J. Anim. Sci.* 65(1):259.

Park, B. S. 1991. Non–Invasive, Objective Measurement of Intramuscular Fat in Beef Through Ultrasonic A–Mode and Frequency Analysis. Phd Dissertation, Texas A&M University College Station, Texas 77843-2117

Shearer, S. A. and R. G. Holmes. 1990. Plant identification using color co–occurrence matrices. *Transactions of the ASAE* 33(6):2037–2044.

Smith, D. and P. Stanford. 1990. A random walk in hamming space. In *Proc. Int. Joint Conf. on Neural Networks,* 2:465–470 San Diego, CA: IEEE.

Thane, B. R. 1991. A neural network application to estimate marbling in ultrasonic B–mode images of beef cattle. Tech. Report AGEN–685. College Station, TX: Dept. of Agricultural Engineering, Texas A&M University.

Thane, B. R. 1992. Prediction of intramuscular fat in live and slaughtered beef animals through processing of ultrasonic images. Thesis, Texas A&M University College Station, Texas 77843-2117.

Thompson, J. 1991. Use of X–ray CT to predict intramuscular fat content in beef. Internal Comm., Dept. of Animal Sci., Univ. of New England, Australia.

USDA 1980. *Official U.S. standards for grades of carcass beef.* Washington, DC: Food Safety and Quality Service.

Wassenberg, R. L., D. M. Allen, and K. E. Kemp. 1986. Video image analysis prediction of total kilograms and percent primal lean and fat yield of beef carcasses. *Journal of Animal Science* 62:1609–1616.

Whittaker, A. D., B. S. Park, J. D. McCauley, and Y. Huang. 1991. Ultrasonic signal classification for beef quality grading through neural networks. In *Proceedings of the 1991 Symposium on Automated Agriculture for the 21st Century*, 116–125. St. Joseph, MI: ASAE.

Shorter version submitted to Transactions of ASAE.