

NONDESTRUCTIVE EVALUTAION
OF BEEF PALATABILITY

By

JEYAMKONDAN SUBBIAH

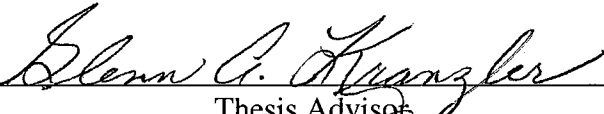
Bachelor of Engineering
Tamil Nadu Agricultural University
Kumalur, India
1997

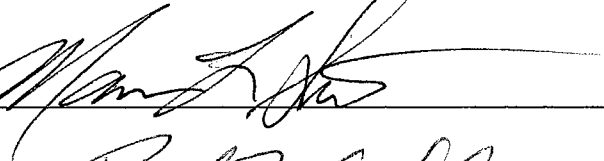
Master of Science
University of Manitoba
Winnipeg, Canada
1999

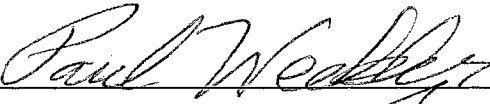
Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirement for
the Degree of
DOCTOR OF PHILOSOPHY
May, 2004

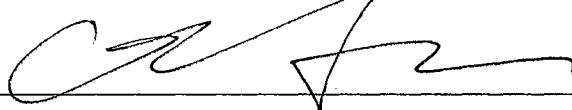
NONDESTRUCTIVE EVALUTAION
OF BEEF PALATABILITY

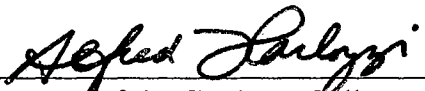
Thesis Approved:


Thesis Advisor








Dean of the Graduate College

ACKNOWLEDGEMENTS

It gives me a great pleasure to present my gratitude to my advisor Dr. Glenn Kranzler for providing advice and guidance throughout the study. While providing independence to carry out this research work, his patience and attention to details kept me focused. Even on occasions when things were not going smoothly, he was always there to support me. His encouragement and support allowed me to participate in several professional development activities. I would like to express my sincere appreciation for his interest in my career and personal life. I am fortunate to have him as my mentor.

I would like to thank my Advisory Committee members, Dr. Marvin Stone, Dr. Paul Weckler, and Dr. Guoliang Fan for their guidance. My course project on near-infrared spectroscopy in Dr. Stone's Instrumentation Course provided the foundation for one of the three major components of my dissertation work. Dr. Weckler presented the idea of investigating X-ray imaging during my research proposal, and it became the third component of my dissertation. Dr. Fan's courses on Digital Signal Processing, Computer Vision, and Wavelets provided fundamentals and basics for conducting the research on computer vision, which is the first major component of my dissertation.

Thanks to Dr. Glenn Kranzler and Dr. Ron Noyes for support as my supervisors in my position of Research Engineer. This position provided rich experience in grant proposal writing and also provided financial support to complete my study. Thanks to Biosystems Engineering Department for funding my position.

My sincere thanks to Dr. Brad Morgan for providing Meat Science expertise in my dissertation work. He always had faith in my work and provided motivation throughout the study. Several of his graduate students, including Sarah Rust, Laura Locke, Aaron Elam, Chad Carr, and Keith Charmasson collected beef samples and conducted shear-force analysis. Dr. Morgan's contacts with beef industry enabled me to collect spectral readings in beef packing plants. He generated funds from the National Cattlemen's Beef Association and the Oklahoma Beef Industry Council to conduct the NIR portion of this study. In addition, grants supporting my research came from the Food and Agricultural Products Research & Technology Research Initiative Program (2000, 2001, 2002). Appreciation is expressed to Jake Nelson, FAPC Meat Processing Manager, for providing assistance in sample preparation. Sincere thanks go to Dr. Tommy Wheeler and personnel at the U.S. Meat Animal Research Center for generating slice shear-force values for validating the NIR portion of this study.

Thanks to Nisha Biju whose thesis work provided a jumpstart on computer vision research. I learned several image processing techniques by working with Nilanjan Ray and Dr. Scott Acton on image segmentation. Thanks to Anand Lakshmikanth for helping in image acquisition. I would like to express my gratitude to Nachiket Kotwaliwale, who developed the X-ray imaging system used in this study.

Thanks to my family members for supporting my pursuit of an advanced degree abroad. Special thanks to my wife, Latha, for her persistent support enabling me to complete this study on time. It is a pure joy to see my son, Vishal, growing day by day. He has added wonder to my life!

To my Family

TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION	1
Objectives	3
Dissertation Organization	3
References.....	5
II. REVIEW OF LITERATURE.....	8
Current Beef Grading System.....	8
Objective Grading.....	9
Ultrasound.....	10
Magnetic Resonance Imaging (MRI).....	11
Video Image Analysis.....	11
Kansas State University	12
Danish Beef Classification Center	12
Australian VIAscan®	14
Canadian Vision System	17
MARC Beef Carcass Image Analysis System	22
Other Studies	22
Tenderness	24
Economic Importance of Beef Tenderness	25
Biological Aspects of Tenderness.....	26

Chapter	Page
Biochemical Aspects of Tenderness	27
Tenderness Standards.....	27
Sensory Analysis	27
Warner-Bratzler Shear Force.....	27
Slice Shear Force.....	28
Methods for Predicting Tenderness	30
Tenderness Probes.....	30
Colorimeter.....	31
BeefCam®	31
Slice Shear Force.....	32
Image Texture	34
Near-infrared Spectroscopy	35
Pioneering Work	36
Matforsk Norwegian Research Group.....	36
Teagasc Irish Research Group.....	38
Other Work.....	39
X-ray Absorption	40
References.....	41
III. COMPUTER VISION SEGMENTATION OF THE <i>LONGISSIMUS DORSI</i> FOR BEEF QUALITY GRADING	55
Abstract.....	55
Introduction.....	56
Materials and Methods.....	59

Chapter	Page
Samples	59
Hardware	60
Reference Segmentation.....	60
Image Processing	60
Fuzzy C-means Clustering Algorithm.....	61
Removal of Intermuscular Fat.....	64
Morphological Operations.....	64
Segmentation Accuracy	67
Classification Error	67
Average Error Pixel Distance.....	68
Resolution and Accuracy	69
Results and Discussion	70
Conclusion	73
Acknowledgements.....	74
References.....	75
IV. BIOEQUIVALENCE ANALYSIS FOR BEEF QUALITY GRADING USING COMPUTER VISION.....	78
Abstract.....	78
Introduction.....	79
Confidence Interval Method.....	81
Selection of 'c'	82
Beef Quality Grading.....	83
Marbling Score.....	86

Chapter	Page
Color Score	87
Quality Grades	88
Conclusion	88
Acknowledgements.....	89
References.....	89
V. PREDICTING BEEF TENDERNESS FROM STATISTICAL IMAGE TEXTURAL FEATURES.....	92
Abstract	92
Introduction.....	93
Materials & Methods	95
Samples	95
Computer Vision System	96
Tenderness Reference	98
Image Processing	98
Preprocessing for Full-scale Images.....	98
Color Space Conversion.....	99
Statistical Textural Features	101
Marbling Features	104
Statistical Models	104
Results.....	105
Prediction of Tenderness from Full-scale Images.....	105
Prediction of Tenderness from Close-up Images	105
Discussion	106

Chapter	Page
Conclusion	107
Acknowledgements.....	108
References.....	108
VI. TEXTURE ANALYSIS OF BEEF IMAGES USING THE GABOR FILTER	112
Abstract.....	112
Introduction.....	113
Materials and Methods.....	114
Samples	114
Computer Vision System	115
Tenderness Reference Measurement.....	115
Image Texture	115
Digital Filter	116
Gabor Filter.....	118
Implementation of Gabor Filter.....	126
Canonical Discriminant Analysis.....	129
Results.....	130
Conclusion	134
Acknowledgements.....	134
References.....	135
VII. PREDICTING BEEF TENDERNESS FROM WAVELET TEXTURAL FEATURES	138
Abstract.....	138

Chapter	Page
Introduction.....	139
Materials and Methods.....	142
Samples	142
Computer Vision System	143
Tenderness Reference Measurement.....	143
Image Texture	143
Need for New Transform	144
Short-time Fourier Transform.....	144
Wavelet Transform.....	145
Continuous Wavelet Transform	146
Mother Wavelet.....	147
Discrete Wavelet Transform	148
Filter-bank Representation of DWT.....	148
Wavelet Decomposition of Images	151
Wavelet Features	153
Color Space Conversion.....	156
Statistical Evaluation of the System.....	156
Results & Discussion	158
Wavelet Textural Features	158
Regression Model	158
Evaluation of the System	160
Conclusion	162
Acknowledgements.....	162
References.....	163

Chapter	Page
VIII. PREDICTING BEEF TENDERNESS USING NEAR-INFRARED SPECTROSCOPY	167
Abstract	167
Introduction	168
USDA Beef Quality Grading	168
Instrument Grading	169
Near-infrared Spectroscopy	171
Pioneering Work	171
Matforsk Norwegian Research Group.....	172
Teagasc Irish Research Group.....	173
Other Work.....	174
Objectives.....	176
Materials & Methods	176
Meat Samples	176
NIR Spectrometer.....	178
Reflectance Measurement	178
Shear-Force Measurement	179
Model Development.....	180
Evaluation of Statistical Model.....	181
Results & Discussion	182
Meat Samples	182
NIR Readings.....	182
Conclusions.....	187
Acknowledgements.....	187

Chapter	Page
References.....	188
IX. PREDICTING BEEF TENDERNESS USING X-RAY ATTENUATION PROPERTIES.....	192
Abstract.....	192
Introduction.....	193
Materials & Methods	194
Principles of X-rays.....	194
X-ray Imaging System	196
Samples	198
X-ray Imaging	198
Slice Shear-Force Measurement	199
X-ray Data Analysis	200
Results & Discussion	202
Conclusions.....	204
References.....	205
X. CONCLUSIONS	208
Summary.....	208
Suggestions for Future Research	214
COMPLETE REFERENCE LIST	219
APPENDIX A: DEVELOPMENT AND CALIBRATION OF A SOFT X-RAY DIGITAL IMAGING SYSTEM FOR AGRICULTURAL PRODUCTS.....	241
Abstract.....	241
Introduction.....	242

Chapter	Page
Principle of X-ray Imaging	243
Materials and Methods.....	245
Equipment	245
X-ray Tube	246
X-ray Camera	246
System Calibration and Testing	247
Dark Current.....	248
Flat-Field Correction	248
Effect of Equipment Variables	248
Repeatability of System	249
Development of Model.....	250
Target Mass Attenuation Coefficient	252
Results and Discussion	252
Effect of X-ray Tube Current.....	252
Effect of X-ray Tube Voltage	254
Effect of Integration Time.....	255
Imprecision.....	257
Statistical Model	257
Evaluation of Statistical Models	258
Target Attenuation Coefficient	259
Conclusions.....	260
Acknowledgements.....	261
References.....	262

LIST OF TABLES

Table	Page
2.1. Evaluation of VIAscan® for predicting saleable meat yield (Cannel et al., 1999) ...	16
2.2. Percentage of carcasses that fall within one subclass for the prediction in the validation set (Allen and Finnerty, 2000)	17
2.3. Accuracy (R^2) and repeatability (standard deviation, SD) of CVS for predicting LMA (Steiner et al., 2003b).....	21
3.1. Evaluation of segmentation accuracy and computation time	72
4.1. Numerical values for marbling scores	84
4.2. Numerical scale for quality grades	85
4.3. Equivalence analysis for color and marbling scores and quality grades	87
5.1. Descriptive statistics of carcass data.....	96
5.2. Confusion matrix for tenderness classification, full-scale images.....	105
5.3. Confusion matrix for tenderness classification, close-up images.....	106
6.1. Distance matrix between categories in canonical feature space	132
6.2. Resubstitution results of canonical discriminant model	132
6.3. Cross-validation results.....	133
7.1. Descriptive statistics from sample carcass data.....	142
7.2. Wavelet textural feature selected by the regression model.....	159
7.3. Statistical differences of the mean observed shear-force values between “certified tender” and “not certified tender” groups	160
8.1. Carcass and shear data summary of all samples	182

Table	Page
8.2. Mean shear-force values of Certified and Not Certified Tender groups for Phase II (calibration) samples.....	185
8.3. Mean shear-force values of Certified and Not Certified Tender groups for Phase III (validation) samples.....	186

LIST OF FIGURES

Figure	Page
2.1. Ribeye sectioned between 12 th and 13 th ribs.....	8
2.2. Danish BCC-2 system. Courtesy: SFK Technology A/S, Denmark.....	14
2.3. Linear measurements made from whole carcass image, CVS system. Courtesy: Research Management Systems, USA Inc., Fort Collins, CO.....	18
2.4. CVS captures ribeye image. Courtesy: Research Management Systems, USA Inc., Fort Collins, CO.....	19
2.5. Calculation of ribeye area, subcutaneous fat thickness, and marbling from ribeye image, CVS system. Courtesy: Research Management Systems, USA Inc., Fort Collins, CO	20
2.6. Sample slice acquisition equipment.....	29
2.7. Sample slice from ribeye	29
2.8. Comparison of SSF, Colorimeter, and BeefCam® (Wheeler et al., 2002).....	34
3.1. Segmentation of longissimus dorsi (l.d.) muscle by adaptive erosion and dilation using convex hull fitting. White line surrounding the muscle is the convex hull....	58
3.2. Green-band histogram of beef steak image	61
3.3. Jarvis's march algorithm to determine convex hull. Vertices of convex hull determined by the algorithm are indicated by cross marks.....	66
3.4. Mismatch between the segmented image (Fig. 1f) and the reference image. White regions indicate mismatch, whereas black regions indicate match.....	69
5.1. Video image analysis system.....	97
5.2. Full image of ribeye steak, 8-mm lens.....	97

Figure	Page
5.3. Close-up image from central ribeye, 50-mm lens.....	98
6.1. Gabor filter design (Adapted from Manjunath and Ma, 1996).....	119
6.2. Impulse and frequency response of Gabor filter ($K=3; S=3; n=0; m=0,1,2$)	121
6.3. Frequency response of Gabor filter ($K=3; S=3; n=0; m=0$)	123
6.4. Frequency response of Gabor filter ($K=3; S=3; n=0; m=1$)	123
6.5. Frequency response of Gabor filter ($K=3; S=3; n=0; m=2$)	124
6.6. Impulse and frequency response of Gabor filter ($K=3; S=3; n=1; m=0,1,2$)	125
6.7. Impulse and frequency response of Gabor filter ($K=3; S=3; n=2; m=0,1,2$)	125
6.9. Illustration of a Gabor filter ($S=4, K=6, m=1; n=0$).....	127
7.1. Fourth-order Daubechies (db4) mother wavelet.....	148
7.2. Low- and high-pass filters for implementation of the 4 th order Daubechies wavelet	149
7.3. Wavelet decomposition of images using filters. Downward arrows indicate down- sampling by a factor of 2 along rows or columns. A - approximation, H - horizontal details, V - vertical details, D - diagonal details	151
7.4. Level 1 decomposition of image using the Haar wavelet. A - approximation, H - horizontal details, V - vertical details, D - diagonal details.....	152
7.5. Five-level decomposition of beef images using the db4 wavelet	154
7.6. Error of the computer vision system at various certification levels.....	161
8.1. Scanning a sample in a packing plant using a portable spectrometer with contact probe	178
8.2. Typical absorbance spectrum from beef ribeye	183
9.1. Schematic of X-ray imaging system.....	197
9.2. X-ray image of a beef steak	202

Figure	Page
9.3. R-value image of a beef steak.....	203
A.1. A typical spectrum of X-rays produced by tungsten anode at 40 kV peak voltage with 125 μm Be filter (Siemens, 2001).....	244
A.2. Schematic of the equipment setup	246
A.3. Detector saturation at higher voltages and currents at 460 ms	251
A.4. Blank image (I_0) taken at 50 kVp, 0.22 mA, 460 ms demonstrating detector variation	251
A.5. Effect of current on intensity at lower voltages for blank images	253
A.6. Effect of current on intensity at higher voltages for images with filtered X-ray beam.....	253
A.7. Effect of voltage on intensity of blank images at various currents.....	255
A.8. Effect of integration time on the mean image intensity for eight different combinations of voltage and current.....	256
A.9. Saturation in dark current and signal due to increase in integration time.....	256
A.10. Relationship between voltage and slope of the linear current model.....	258
A.11. Evaluation of model (Eqn. A.9) for calibration images	259
A.12. Attenuation coefficient of polystyrene at different X-ray energies.....	260

CHAPTER I

INTRODUCTION

The beef industry is the largest food and fiber industry in the U.S. (NCBA, 2001). This industry relies on the USDA quality grading system to classify carcasses into groups based on expected meat palatability. Beef grading standards were first developed by the United States Department of Agriculture (USDA) in 1926 to facilitate marketing (USDA, 1997). A trained USDA employee, working independently of both the producer and the packer, determines beef grades. Meat retailers and restaurants use USDA grades as a value guideline.

Typical packing plant production rates allow only 9 to 18 seconds per carcass to assign grades (Belk et al., 1998). Although highly trained, USDA graders are subject to fatigue and emotional strain, which can affect the decision-making process. Cross et al. (1980) reported as much as 21% error in manual grading. Due to the subjective nature of the process, there are inconsistencies in grade assignment. An objective “instrumented” carcass evaluation procedure is needed to ensure consistent quality. The “Value-Based Marketing Task Force” of the Cattlemen’s Beef Board and the “National Beef Instrument Assessment Plan” of the National Live Stock and Meat Board identified video image analysis, or computer vision, as the most promising instrumented technique for industrial application (Belk et al., 1996).

Quality grades are based on marbling levels (abundance and distribution of intramuscular fat within the lean) and on physiological maturity of the carcass. Five

maturity groups have been defined. They are; A, B, C, D, and E, in order of increasing maturity. The National Beef Quality Audit-2000 reported that almost all carcasses processed in the U.S. are of the younger “A” maturity (McKenna et al., 2002). As a result, USDA quality grades rely heavily on marbling as a predictor of palatability. Wheeler et al. (1994) reported that marbling explained, at most, 5% of the variation in palatability traits. The current grading system, then, is not successful in segmenting carcasses on the basis of palatability.

Meat palatability traits include tenderness, juiciness, and flavor of the cooked product. Tenderness alone is addressed in this study, because:

- Consumers hold tenderness as the primary factor in eating satisfaction.
- Consumers can differentiate tenderness level and are willing to pay a premium for tenderness (Boleman et al., 1997).
- When tenderness is guaranteed on the label, consumers will pay more (Lusk et al., 2001).
- Tenderness is positively related to juiciness and flavor (Winger and Hagyard, 1994).
- The tenderness coefficient of variability found in U.S. carcasses is almost twice that of juiciness and flavor (Koochmaraie, 1995). Predicting a meat property that exhibits larger variability is more easily instrumented.

The beef industry is looking for an instrument that can take readings on fresh beef to predict tenderness of the cooked product. It is common industry practice to age beef carcasses for 10-14 days to enhance tenderness. Increase in tenderness during aging varies considerably from animal to animal (Wheeler and Koochmaraie, 1994). The

variability effects of postmortem aging and cooking on ultimate tenderness present an immense challenge in the development of a predicting instrument. As a result, direct measurement of beef tenderness is not now included in carcass grading. Because carcasses are not priced on tenderness, producers lack incentive to produce tender carcasses. Thus, consumer preference is not transmitted to the producer.

Objectives

The overall objective of this study is to develop nondestructive methods to predict beef palatability. Initial effort was directed toward developing a method to predict USDA quality grade. Acknowledging the limitation of USDA quality grading in addressing consumer preference, this investigation subsequently examined three nondestructive methods to predict aged, cooked-beef tenderness:

1. Computer vision,
2. Near-infrared spectroscopy,
3. X-ray imaging.

Dissertation Organization

Chapter 2 reviews pertinent literature related to this study. Chapters 3-9 are prepared as stand-alone chapters compatible with the American Society of Agricultural Engineers (ASAE) Transactions format. In Chapter 10, results of Chapters 3-9 are summarized, and suggestions for future work are presented. A synopsis of Chapters 3-9 is presented below.

Chapters 3 and 4 describe the development of a computer vision system to predict USDA quality grade. USDA inspectors assign quality grades from visual appraisal of the

longissimus dorsi (l.d.) muscle from the carcass sectioned between the 12th and 13th ribs. Segmentation of the l.d. muscle from the steak image is the primary step in developing a successful computer vision system for grading beef. In Chapter 3, an adaptive segmentation algorithm is developed to separate the l.d. muscle from the beef steak image. Features are extracted from the segmented l.d. muscle and used to predict USDA quality grade. A bioequivalence statistical analysis method is implemented in Chapter 4 to evaluate if grades predicted by the computer vision system are *significantly equivalent* to grades assigned by expert graders.

Chapters 5, 6, and 7 describe the development of the computer vision system to predict tenderness. Chapter 5 uses the gray-level sum-and-difference histogram to extract statistical textural features. Chapters 6 and 7 describe multiresolution textural analysis using Gabor filters and wavelets, respectively.

Development of an NIR spectroscopy system is described in Chapter 8. An NIR spectrometer with reflectance probe is used to acquire online spectral scans from beef carcasses. A chemometric model is developed to predict cooked-beef tenderness. The system was tested online in two Texas packing plants.

Chapter 9 examines the feasibility of using beef X-ray attenuation properties to predict tenderness. Appendix A describes the development and calibration of an X-ray imaging system to enable measurement of attenuation properties of biological materials.

References

- Belk, K. E., J. A. Scanga, J. D. Tatum, J. W. Wise, and G. C. Smith. 1998. Simulated instrument augmentation of USDA yield grade application to beef carcasses. *Journal of Animal Science* 76(2): 522-527.
- Belk, K. E., J. D. Tatum, G. Dolezal, B. Morgan, and G. C. Smith. 1996. Meat composition measurement: status of applied research on instrument assessment of composition since completion of the 1994 National Beef Instrument Assessment Planning Symposium. In *Reciprocal Meat Conference Proceedings*, 49: 172-174.
- Boleman, S. J., S. L. Boleman, R. K. Miller, J. F. Taylor, H. R. Cross, T. L. Wheeler, M. Koohmaraie, S. D. Shackelford, M. F. Miller, R. L. West, D. D. Johnson, and J. W. Savell. 1997. Consumer evaluation of beef of known categories of tenderness. *Journal of Animal Science* 75(6): 1521-1524.
- 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. Final report to office of inspector general, USDA.
- Koohmaraie, M. 1995. The biological basis of meat tenderness and potential genetic approaches for its control and prediction. Clay Center, NE.: USDA-ARS, Roman L. Hruska U.S. Meat Animal Research Center. Available at: http://meats.marc.usda.gov/MRU_WWW/ICMST95/ICMST95.html. Accessed on 9 February 2004.

- Lusk, J. L., J.A. Fox, T.C. Schroeder, J. Mintert, and M. Koohmaraie. 2001. In-store valuation of steak tenderness. *American Journal of Agricultural Economics* 83(3): 539-550.
- McKenna, D. R., D. L. Roebert, P. K. Bates, T. B. Schmidt, D. S. Hale, D. B. Griffin, J. W. Savell, J. C. Brooks, J. B. Morgan, T. H. Montgomery, K. E. Belk, and G. C. Smith. 2002. National Beef Quality Audit-2000: survey of targeted cattle and carcass characteristics related to quality, quantity, and value of fed steers and heifers. *Journal of Animal Science* 80(5): 1212-1222.
- NCBA. 2001. The U.S. beef industry: its impact on the American economy. Fact Sheet. Centennial, CO.: National Cattlemen's Beef Association. Available at: http://www.beef.org/dsp/dsp_content.cfm?locationId=710&contentType=1&contentId=252. Accessed on 11 February 2004.
- USDA. 1997. United States standards for grades of carcass beef. Effective date January 31, 1997. Washington, DC.: Livestock and Seed Division of the Agricultural Marketing Service. Available at: <http://www.ams.usda.gov/lsg/stand/standards/beef-car.pdf>. Accessed on 9 February 2004.
- Wheeler, T. L., L. V. Cundiff, and R. M. Koch. 1994. Effect of marbling degree on beef palatability in *Bos taurus* and *Bos indicus* cattle. *Journal of Animal Science* 72(12): 3145-3151.
- Wheeler, T. L., and M. Koohmaraie. 1994. Prerigor and postrigor changes in tenderness of ovine longissimus muscle. *Journal of Animal Science* 72(5): 1232-1238.

Winger, R. C., and C. J. Hagyard. 1994. Chapter 4: Juiciness - its importance and some contributing factors. In *Quality attributes and their measurement in meat, poultry and fish products. Advances in meat research series*. Vol. 9. 94-124. A. M. Pearson and T. R. Dutson, eds. London, UK.: Blackie Academic & Professional.

CHAPTER II

REVIEW OF LITERATURE

Current Beef Grading System

Since 1927, the Meat Grading and Certification Branch of the USDA has been providing a voluntary beef grading and stamping service to the beef industry. USDA inspectors assign both quality and yield grades. Quality grade categorizes the carcasses based on beef palatability factors, whereas yield grade gives an estimate of saleable red meat. Quality grades are assigned from visual appraisal of the *longissimus dorsi* (l.d.) muscle sectioned between the 12th and 13th ribs (Fig. 2.1).



Figure 2.1. Ribeye sectioned between 12th and 13th ribs.

Quality grades are based on marbling levels (abundance and distribution of intramuscular fat on the surface of the l.d. muscle) and physiological maturity of the carcass. The USDA defines ten marbling scores and has published photographs

illustrating these scores. These photographs serve as references, when needed, for official marbling scores. Five levels of carcass maturity are defined on the basis of color, texture, and firmness of the l.d. muscle and on the level of ossification of cartilage at the ends of the dorsal spinous processes of the vertebral column. USDA yield grades are determined through the assessment of the following characteristics: area of the l.d. between the 12th and 13th rib, hot carcass weight, thickness of fat opposite the l.d., and an estimation of internal kidney, pelvic, and heart fat. USDA yield grades are used to determine the percentage of boneless, retail cuts that can be obtained from the major portions of the carcasses. USDA publishes detailed beef carcass grading procedures (USDA, 1997).

Due to the subjective nature of the grading process, there are inconsistencies in grade assignment. Graders tend to favor mid-range estimates instead of using the whole quality grade range from Prime to Choice to Select to Standard. The National Beef Quality Survey-2000 reports that 91.4% of carcasses grade either USDA Select or Low Choice (McKenna et al., 2002). Subjective grading does not foster confidence in the current marketing system. Lack of effective incentive for producing quality carcasses has led to beef industry loss of protein market-share to the chicken industry (Lusk et al., 2001). An objective “instrumented” carcass evaluation procedure is needed to ensure consistent quality and to facilitate value-based marketing.

Objective Grading

Clearly, the current grading system has some undesirable characteristics. The beef industry has been searching for an objective grading system for more than two decades. In 1979, the FSQS (now the Agricultural Marketing Service and the Food

Safety Inspection Service of the USDA) and the National Aeronautical Space Agency (NASA) jointly funded a study at the Jet Propulsion Lab managed by California Institute of Technology, Pasadena, CA to identify technologies for beef quality grading (Cross and Whittaker, 1992). NASA named ultrasound and video image analysis (VIA) as potential technologies for objective grading.

Ultrasound

Ultrasound is sound with frequencies above the audible range for humans (>20 kHz). Interaction of ultrasound with a sample can be used to determine material properties. Two modes of ultrasound are used. The A-mode (amplitude modulation, one-dimensional representation) measures the amplitude of the sound wave after interacting with the sample. The B-mode (brightness modulation) is a two-dimensional representation of a material. For carcass evaluation, sound waves of 3.5 MHz are commonly used (Perkins et al., 2003).

Due to development by the medical community, there are off-the-shelf, portable, real-time ultrasound instruments available on the market. This technology can be successful in predicting marbling scores of live cattle, rather than carcasses (Cross and Belk, 1994). The major problem with this technique is that the instrument must be held at a specific location on the live animal in order to obtain useable image quality. If technicians are not well-trained, large errors can result. Based on the predicted marbling score in live animals using ultrasound, suitable harvest date for slaughter cattle can be identified. However, ultrasound is not suitable for predicting beef grades (Smith, 1999).

Magnetic Resonance Imaging (MRI)

When a strong magnetic field is applied to a biological material such as a food product, certain atomic nuclei (with spin angular momentum) interact. Orientation of the atomic particles reaches an equilibrium condition. A resonant radio frequency is applied to excite the particles resulting in a non-equilibrium state. The atomic particles then return to a state of equilibrium by releasing decaying energy. This energy is sensed as a magnetic resonance signal (Stroshine, 1998). The hydrogen nucleus has the strongest magnetic moment. Because hydrogen is a major chemical component in food products, MRI has been successfully used in investigating water content and chemical composition.

MRI has been successfully used for investigating water interaction in meat during freezing (Renou et al., 2003) and high-pressure treatment (Bertram et al., 2004). Antequera et al. (2003) obtained accuracy of 83% for predicting intramuscular fat and 70% for sensory and chemical traits of dry-cured loin using MRI. Mitchell et al. (2001a) used MRI for body composition analysis of live pigs. Bonny et al. (2000a, b) demonstrated that MRI images of bovine muscles show features of connective tissue. However, these authors did not use MRI to predict connective tissue content.

Video Image Analysis

Video image analysis systems are suitable for simulating human vision. A VIA system consists of a camera to capture video images of lighted objects. In case of an analog video camera, a frame grabber is used to digitize image signals. A digital image has finite resolution (sampling) and bit-depth (quantization). A gray-scale and a color image are represented by a two-dimensional and a three-dimensional matrix, respectively. Various image processing algorithms (mathematical operations on the image matrix) can

be implemented to segment an object of interest in the image frame and to extract specific features. Pattern recognition methods incorporating statistical algorithms or neural networks are used to predict food quality from extracted features. Key advantages of VIA are: nondestructive, objective, and consistent. Once developed, VIA is fast and can be implemented in an industry environment. In some cases, VIA has even more advantage in that it can capture images beyond the visible range (near-infrared, X-ray, etc.) using suitable sensors. Because of these advantages, VIA systems have been developed around the world for beef carcass grading. Several are explained below:

Kansas State University

After NASA identified VIA as a promising technology for beef quality grading, Kansas State University won the contract from the USDA ARS in 1980 to develop a VIA system. The system acquired and processed gray-scale images of 244 x 248 pixel resolution. A frame positioned the video camera over the ribeye in the hanging carcass. Lighting was provided by four fluorescent lamps. An infrared stop filter and a red filter on the lens enabled the use of simple thresholding to segment background, lean, and fat in the digital image (Lenhert and Gilliland, 1985). The system was tested in the U.S. Meat Animal Research Center, Clay Center, NE from 1981-83. Cross et al. (1983) and Wassenberg et al. (1986) reported that the system had more potential for yield grading than for quality grading.

Danish Beef Classification Center

The Danish Meat Research Institute developed Beef Classification Centre (BCC), a VIA system for predicting yield grade according to the EUROP scheme regulated by European Union (Petersen et al. 1989). The EUROP scheme defines fatness and

conformation (shape) on a scale of 1 to 5. Each class is further divided into three subclasses (Nielsen et al., 1997). In addition to weight and type of animal, these scales are used to determine yield grade of the carcass. The first generation (BCC-1) used a monochrome camera (Signe et al., 1989). The second generation BCC-2 used a 3-CCD color camera (Borggaard et al., 1996). A movable frame grabbed a carcass on the rail and held a predefined position for imaging (Fig. 2.2). A green background screen was used. The lighting system consisted of two large lamps with 18 halogen bulbs. One image was taken with lights ON and another image with lights OFF. The two images were subtracted to eliminate the effect of ambient light. In addition, structured lighting (Fig. 2.2) was used to reconstruct the three-dimensional shape of the carcass (Madsen and Thodberg, 1999). The system was calibrated using a 1 x 2-m calibration plate to correct for geometric distortion and lens distortion. Color and gray-scale calibration plates were imaged for color calibration (Madsen and Thodberg, 1999). After segmenting the carcass from the background using the green/red ratio for thresholding, carcass contour information was collected. The carcass area was divided into 60 regions. Mean, variance, and covariance of the red, green, and blue (RGB) color features were extracted. From the image obtained by structured lighting, the thickness was calculated at 340 points on the carcass. Principal component analysis was performed to reduce dimensionality of features. The BCC-2 gave errors 16-22% lower than those of the classifier (on-line manual grader), when grades assigned by the inspector or expert grader were used as reference. The R^2 value for predicting conformation, fatness, fat color, percent saleable meat, and ribeye area were 0.93, 0.75, 0.89, 0.70, and 0.85, respectively (Borggaard et al., 1996). The system could analyze 80 carcasses per hour.



Figure 2.2. Danish BCC-2 system. Courtesy: SFK Technology A/S, Denmark.

Australian VIAscan[®]

The Australian Meat Research Corporation developed a color VIA system called (VIAscan[®]) to predict beef grades. The VIAscan[®] consists of two vision systems; hot assessment system (HAS), and chiller assessment system (CAS). The HAS acquires a video image of the whole carcass on the slaughter/dressing floor and evaluates the surface contour and the outside surface fatness (Cannel et al., 1999). The CAS system consists of a frame that encloses a solid-state RGB video camera within a stainless steel shroud (Benn et al., 1998). At the bottom of the frame, structural guides ensure proper alignment of the camera with the carcass. The lens is fitted with a polarizing filter. The lamp assembly consists of a finned heat-sink housing a 12-V dichroic lamp fitted with an infrared filter, diffuser, and polarizer. The infrared filter is provided to minimize heat

damage to the meat. Matt reflectors surround the frame to provide uniform illumination of the meat surface. Polarization filters in the lens and lamp assembly are oriented to minimize specular reflection. A computer for image analysis and a monitor to display video images are enclosed in a waterproof cabinet. A frame grabber in the computer digitizes to a 24-bit true color image with 512 x 512 spatial resolution. Both spatial and color calibrations are performed. Images of black, gray, white, and color reference boards are acquired for color calibration. Coefficients of a quadratic equation for relating the observed color values to the reference color values were developed. This equation is subsequently used for color correction (Benn et al., 1998).

The CAS system measures ribeye area, marbling scores, meat color, fat color, and subcutaneous fat thickness. First, the image is globally thresholded in the red band to separate objects from background. Global fixed thresholding is performed on the R/G ratio image to separate fat and lean. The largest object is identified. A heuristic procedure is performed to remove extraneous tissue from the ribeye muscle. Ribeye area is calculated by counting the number of ribeye pixels. A local thresholding method is used to segment marbling in the ribeye. After marbling segmentation, average R, G, and B values of pixels classified as lean and marbling are calculated as a measure of lean color and fat color, respectively (Benn et al., 1998).

Cannell et al. (1999) evaluated VIAscan[®] for predicting saleable meat yield using 240 carcasses. Predictions were compared with yield grades assigned by online graders at packing plant speeds and by expert graders working at a comfortable pace. Because online graders make errors in estimating ribeye area (Belk et al., 1998), Cannel et al.

(1999) also evaluated VIAscan[®] as an augmentation tool for estimating ribeye area for online graders. Results are shown in Table 2.1.

Table 2.1. Evaluation of VIAscan[®] for predicting saleable meat yield (Cannel et al., 1999).

Method	Percent variation accounted in fabrication yields at fat-trim levels of:		
	Commodity-trimmed	Closely trimmed	Very closely trimmed
Expert Grader	51	74	74
Online Grader	37	54	54
VIAscan [®]	46	66	71
VIAscan [®] Augmentation	55	75	75

Cannel et al. (1999) concluded that VIAscan[®] estimated fabricated yields more accurately than online graders, but less accurately than expert graders. However, when VIAscan[®] ribeye area estimates were provided to online graders, the accuracy of cutability prediction was slightly higher than that of expert graders. Goering (1999) reported that the VIAscan[®] did not accurately predict quality grade.

Allen and Finnerty (2000) conducted a comprehensive comparative study of three VIA systems; Danish BCC-2, Australian VIAscan[®], and German VBS2000. Conducted in two trials in an Irish beef packing plant, the study compared prediction of carcass conformation, fat class, and saleable meat yield under European classification standards. In the first trial, images of 7,247 carcasses were captured by the three systems. Reference classification scores for 4,278 carcasses were provided to the manufacturers for calibration. Predictions were evaluated on the remaining 2,969 carcasses. All three systems predicted saleable red meat yield with high accuracy (residual standard deviation

of 1.1-1.2%). All systems predicted conformation class with much higher accuracy than fat class (Table. 2.2). One of the criteria set by the European Union Beef Management Committee for approving a VIA system is that at least 88% of the predictions should be within one subclass of the reference panel for fat class. Because this criterion was not met, none of these systems was approved for beef carcass classification in the Irish beef industry.

Table 2.2. Percentage of carcasses that fall within one subclass for the prediction in the validation set (Allen and Finnerty, 2000).

Class	Trial 1			Trial 2		
	BCC-2	VIAscan [®]	VBS2000	BCC-2	VIAscan [®]	VBS2000
Conformation	92.8	91.0	96.5	97.0	94.2	95.4
Fat	80.4	72.0	74.6	79.6	76.1	74.4

Canadian Vision System

Tong et al. (1998) developed a color VIA beef grading system known as Canadian Vision System (CVS). Similar to VIAscan[®], the CVS is also comprised of two subsystems; a carcass image processing subsystem, and a ribeye image processing subsystem. A color calibration matrix is derived by comparing the acquired images of five color boards (black, white, red, green, and blue) with images of the same color reference boards acquired under controlled lighting conditions. Spatial calibration is established by imaging a latticed white board having a grid of black lines spaced at known intervals. A blue, non-reflective background is used for carcass image analysis, because the blue hue has highest contrast with the reddish/yellowish carcass. Simple thresholding is used to detect the contours of the whole carcass. On the carcass contour, anatomical features

such as the tail, lower hip, etc. are identified. Several distance measurements based on anatomical features are made on the carcass contour (Fig. 2.3). These readings are used in a regression equation to predict saleable meat yield.

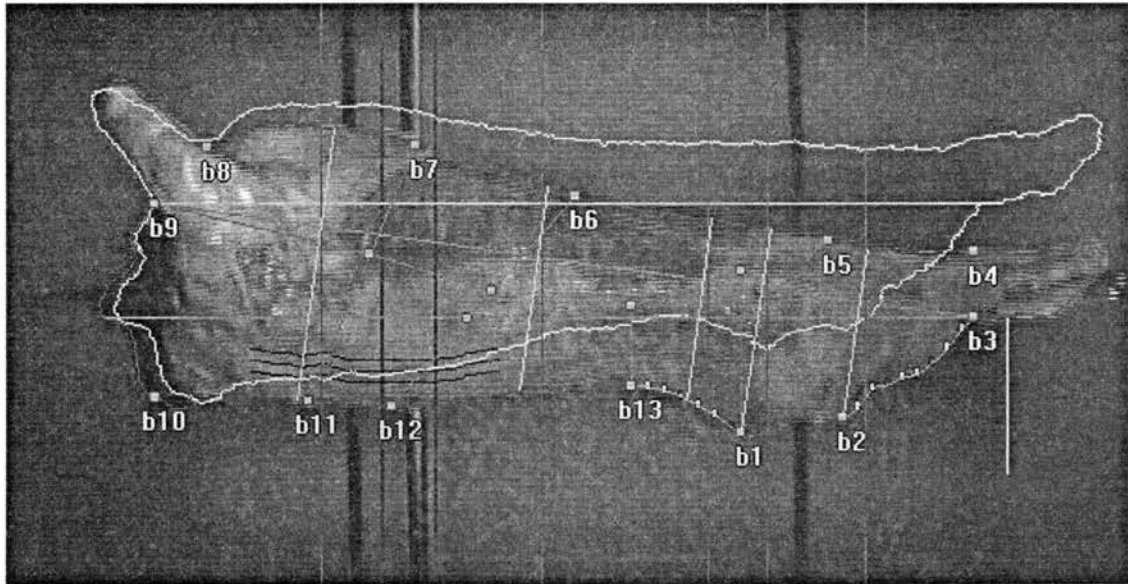


Figure 2.3. Linear measurements made from whole carcass image, CVS system. Courtesy: Research Management Systems, USA Inc., Fort Collins, CO.

For the ribeye image analysis system, an 18% gray card and a Kodak color calibration card are imaged to calibrate the camera for brightness and color, respectively. Figure 2.4 shows the CVS ribeye image analysis system capturing the image of a ribeye. After acquiring the ribeye image, it is corrected for brightness and color. Image contrast is enhanced and thresholding is performed in the brightness and saturation bands to classify fat, lean, and cartilage. The image is then down-sampled by a factor of 8 and 8 along rows and columns, respectively. Using four-connected blob analysis, the largest component of the image is identified. A heuristic approach is implemented to remove extraneous tissue in order to segment the ribeye. Accuracy of the segmentation was not reported. Figure 2.5 shows the segmented l.d. muscle boundary in the ribeye image.

Ribeye area and subcutaneous fat thickness at four points are measured. After identifying marbling and classifying fat flecks into two classes based on size, percent lean area is calculated. The R^2 value of saleable yield prediction was 0.84, based on 20 carcass and 5 ribeye measurements (Tong et al. 1998).

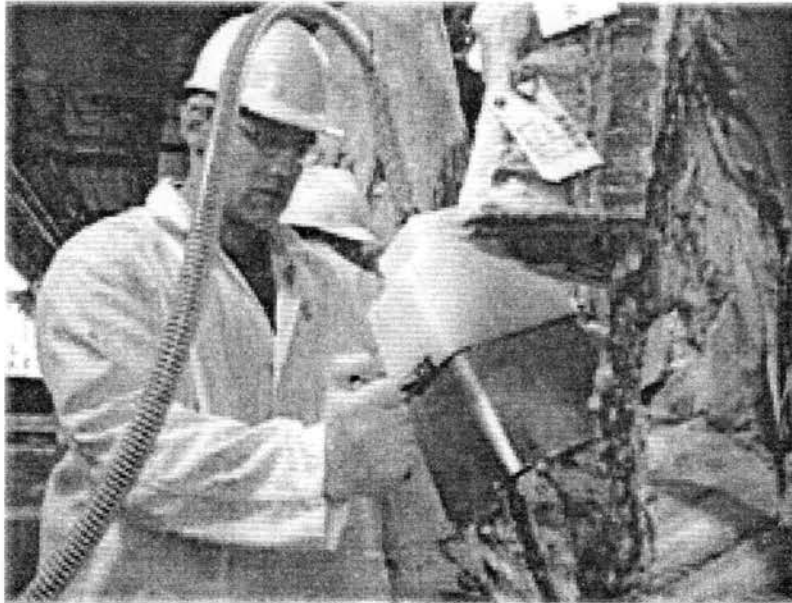


Figure 2.4. CVS captures ribeye image. Courtesy: Research Management Systems, USA Inc., Fort Collins, CO.

Cannell et al. (2002) evaluated CVS for predicting fabricated yields of closely trimmed subprimals using 296 carcasses. CVS accounted for 64% of variation in fabricated yields. When CVS measured ribeye area was provided to expert graders, the variation accounted for increased slightly to 65%. Expert graders and online graders accounted for 67% and 39% of variation in fabricated yields, respectively. This result shows that USDA expert yield grading provides a reasonable accurate estimate (67%) of fabricated yield. However, online graders' estimate of fabricated yield grade was poor (39%) due to the time constraint imposed.

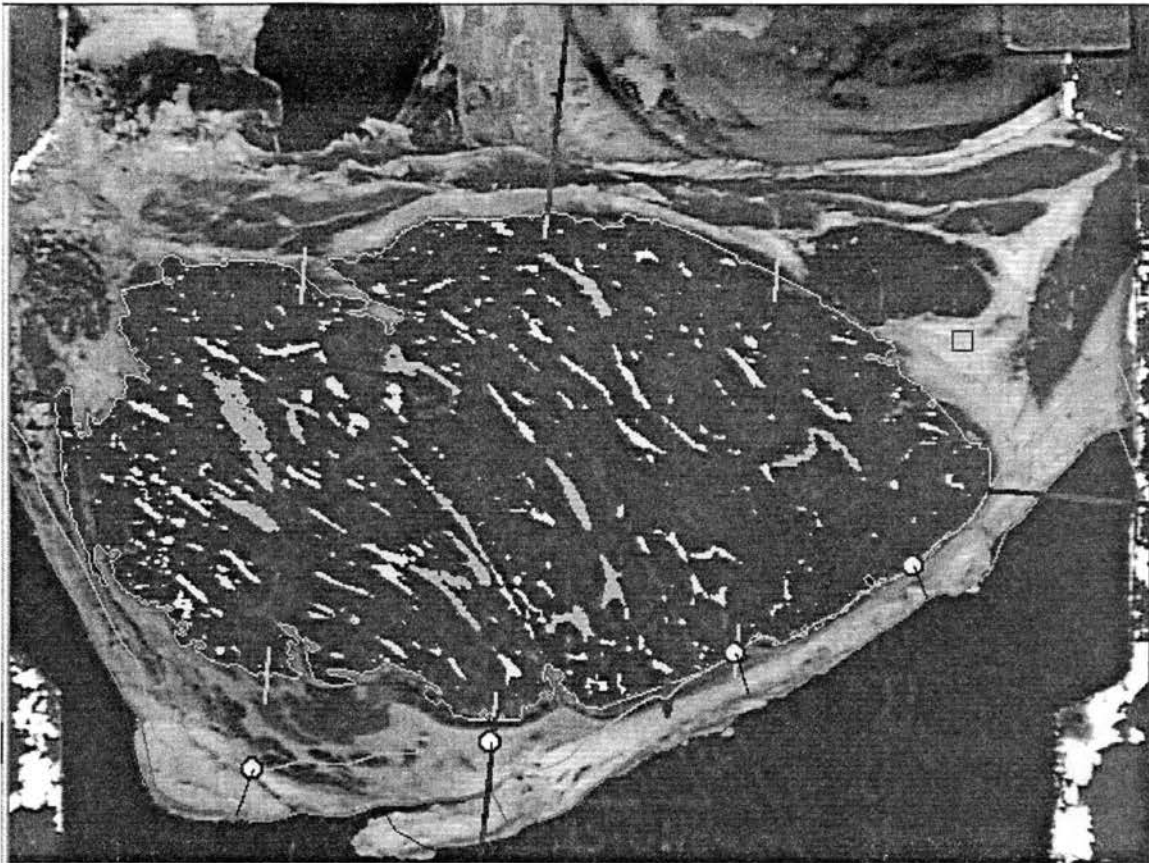


Figure 2.5. Calculation of ribeye area, subcutaneous fat thickness, and marbling from ribeye image, CVS system. Courtesy: Research Management Systems, USA Inc., Fort Collins, CO.

Steiner et al. (2003a) evaluated VIAscan[®] and CVS as augmentation tools for online graders predicting yield grade. In Phase I (505 carcasses), yield grades assigned by online graders provided with ribeye area measurements made by VIAscan[®] and CVS correlated with those assigned by expert graders at 0.93 and 0.95, respectively. In Phase II (290 carcasses), online yield grades augmented by VIAscan[®] and CVS accounted for 60 and 63% of variation in fabricated yields, respectively. In comparison, online yield grades and expert yield grades accounted for 55 and 71% of variation. Steiner et al. (2003a) concluded that augmentation by video image analysis systems improved the accuracy of prediction of fabricated yields.

Steiner et al. (2003b) evaluated accuracy and repeatability of *longissimus dorsi* (ribeye) muscle area (LMA) measured by CVS. Two reference LMA values were used. One was measured by expert graders using a grid scale, and the other by tracing l.d. on acetate paper and subsequently measuring the area using a polar planimeter. The authors tested two CVS instruments to evaluate accuracy and repeatability of LMA measurement by three procedures. In the first procedure, three sequences of images of ribeye were obtained with the camera head in place. In the second procedure, sequential images were acquired with the camera head removed and placed over ribeye between each reading. Carcasses were stationary in the above two procedures. In the last procedure, images were acquired three separate times as carcasses advanced at a chain speed of 360 carcasses/h. Both sides of 50 carcasses were used in this study. CVS predicted traced muscle LMA more accurately than expert-gridded LMA.

Table 2.3. Accuracy (R^2) and repeatability (standard deviation, SD) of CVS for predicting LMA (Steiner et al., 2003b)

Procedure	Acetate/Planimeter traced LMA		USDA requirement	
	R^2	SD (cm ²)	R^2	SD (cm ²)
1	0.94	0.77	>0.95	<1.59
2	0.93	1.10	>0.95	<1.59
3	0.86	2.13	>0.90	<1.90

As expected, CVS predicted LMA more accurately when carcasses were stationary (Table 2.3). USDA (2001) laid out the requirements for approval of VIA for predicting LMA for augmenting yield grades. Repeatability of CVS exceeded the requirements, and the accuracy approached the requirements for stationary LMA measurements (procedures

1 and 2). However, CVS did not clearly meet the requirements for in-motion LMA measurements (Table 2.3).

MARC Beef Carcass Image Analysis System

The Roman L. Hruska U.S. Meat Animal Research Center (MARC) has developed a beef carcass image analysis system. Shackelford et al. (1998) developed an offline imaging system to predict retail product yield from ribeye steaks. Image processing was performed using ImagePro[®] software (Media Cybernetics, Carlsbad, CA). A regression model was developed using 33 samples and was validated with another 33 samples. The R^2 value for predicting retail product yield was 0.91 for the validation data set. Based on this work, MARC collaborated with a major U.S. beef packing plant to develop this technology for online grading of carcasses. Shackelford et al. (2003) evaluated the MARC system using 800 carcasses. Half of the samples were used to develop regression models and the other half were used for validation. The prediction equations accounted for 90% of variation in yield grades assigned by the expert graders, whereas the official yield grades assigned by the online graders accounted for only 73%. The authors reported that the system did not predict quality grade accurately.

Other Studies

Gerrard et al. (1996) utilized image processing techniques to predict color and marbling scores of 60 steaks, offline. After segmenting the l.d. muscle, various color and marbling features were extracted. Stepwise backward elimination regression models were built for predicting color and marbling scores. Reported R^2 values for prediction of color and marbling scores were 0.86 and 0.84, respectively. Their investigation did not predict quality grade.

Lu and Tan (2004) showed that features extracted from the ribeye image predicted lean percentage yield with R^2 value of 0.55. In comparison, the corresponding R^2 value achieved by yield grade factors (fat thickness; percentage of kidney, pelvic, and heart fat (KPH); ribeye area; and hot carcass weight) was 0.76. Their study showed that USDA yield grade factors were more accurate in predicting lean percentage yield.

Hatem et al. (2003) developed a system to predict maturity grade from vertebra images taken from 110 carcasses. First, they manually segmented cartilage and the connected bone from each image. They then extracted features from the manually segmented image to predict maturity grade. Using cross-validation, their neural network model classified carcasses into five maturity grades with an accuracy of 45%. In order to develop an algorithm to automatically segment cartilage and connected bone from the vertebra images, Hatem and Tan (2003) captured the images of the thoracic vertebrae of 83 carcasses. They reported that their algorithm gave satisfactory segmentation results on 69 images.

In summary, there are commercial VIA systems available for predicting yield grades, accurately. USDA yield grade is calculated from four factors: hot carcass weight, KPH, subcutaneous fat thickness, and ribeye area. The first two factors can be objectively measured by weight scale. USDA online graders accurately measure fat thickness; however they make subjective errors in estimation of ribeye area. Belk et al. (1998) envisioned an instrument-augmented grading system in which VIA-predicted ribeye area was combined with online grader assessment of fat thickness and objective measurement of hot carcass weight and KPH to predict yield grade. Recently, USDA approved the use of VIA systems for grader-assisted determination of beef yield grades

(Ishmael, 2000; USDA, 2001). Commercial VIA systems segment l.d. muscle to calculate ribeye area and segment marbling to predict percent lean. It should be noted that no studies have been conducted to test commercial VIA systems for predicting marbling score or quality grade. Therefore, the USDA has not approved VIA systems for predicting quality grade. It can be inferred that none of the systems can estimate USDA quality grade with enough accuracy to be implemented online in packing plants. Nevertheless, these studies show that VIA has great potential for estimating beef quality grades.

Tenderness

Meat palatability is defined as the tenderness, juiciness, and flavor of the cooked product. Of these characteristics, tenderness is the primary determinant of customer satisfaction. Moreover, tenderness is positively correlated with meat juiciness and taste (Winger and Hagyard, 1994). In the 1995 National Beef Quality Audit (Boleman et al., 1998), inadequate beef tenderness was identified as the second largest quality concern of beef among purveyors, retailers, and restaurants. Currently, the USDA grading system does not incorporate a direct measure of tenderness, because there is no accurate, rapid, non-destructive method for predicting tenderness available to beef industry.

Marbling is often considered to be a reliable predictor of meat tenderness. However, the literature shows contradictory results regarding the influence of marbling on tenderness (Jeremiah, 1996). Research has shown that marbling adequately predicts meat tenderness (in young animals) when a wide range of marbling exists within the group being tested. However, the National Beef Quality Audit-1995 revealed that 82%

of carcasses were graded USDA Select or Low Choice, and 83.5% of those carcasses carried marbling scores of either “slight” or “small” (Boleman et al., 1998). Thus, most carcasses graded in U.S. packing plants fall within a small range of marbling and quality grades. Moreover, the National Beef Tenderness Survey (Morgan et al., 1991) showed that consumers experience undesirable toughness in one of every four steaks consumed. It follows that the current quality grading system is not adequately sorting carcasses on the basis of tenderness.

Economic Importance of Beef Tenderness

Boleman et al. (1997) color-coded steaks based on tenderness categories (tender, medium, and tough) and delivered two steaks from each category to 42 families. After cooking, the consumers were asked to complete a survey. The survey showed that consumers were able to detect differences among tenderness categories. Then, the authors posted a \$1.10/kg price difference between each tenderness category. Consumers were asked to purchase steaks. As expected, 94.6% of consumers bought tender steaks, whereas 3.6% and 1.8% of consumers bought medium and tough steaks, respectively. The study showed that consumers can discern tenderness and are willing to pay a premium for tenderness.

Lusk et al. (2001) conducted a consumer survey on perception of tenderness and examined the consumer’s willingness to pay for tenderness in a grocery store setting. They color-coded tender steaks (red) and tough steaks (blue) and provided samples from each category. Sixty-nine percent of consumers showed preference for a tender steak. All customers were provided a free blue-coded (tough) steak for participating in the survey. If they preferred a red-coded (tender) steak, they could bid to exchange it for

their tough steak. If their bid exceeded a predetermined level, they would make the exchange at that bid price. Participants were willing to pay \$2.72/kg for a tender steak, when relying on taste alone. In another treatment, the steaks were labeled as tender and tough instead of being color-coded. Then, 84% of consumers preferred tender steaks. More than half of the participants were willing to pay \$4.06/kg more for a “guaranteed tender” steak.

Shackelford et al. (2001) showed that more than half of consumers are willing to pay \$1.10/kg premium for “guaranteed tender” U.S. Select grade steaks. Note that Select grade has less marbling and is traditionally less preferred by customers.

These studies show that most consumers can discern tenderness and a considerable portion of those consumers are willing to pay a premium for steaks that are “guaranteed tender.”

Biological Aspects of Tenderness

Koohmaraie (1995) states that genetics control about 30% of the variation in beef tenderness within a single breed. It follows that 70% of variation in tenderness is affected by environmental factors. Based on traditional animal breeding theory, it takes about 40 years to improve the shear-force tenderness by 1 kg through progeny testing and breeding programs. The author argues that it is not practical to improve tenderness by genetic means alone due to the time and expense involved.

Biochemical Aspects of Tenderness

Immediately after slaughter, most meat is tender. During rigor development, increase in sarcomere length leads to meat toughness. Meat achieves its maximum toughness between 12 and 24 h postmortem. During this time, an opposite phenomenon called tenderization begins and continues for 14-21 days in meat stored at refrigerated temperatures. This process is also commonly known as “aging.” Proteolytic enzymes break myofibrillar proteins and weaken structural integrity. The result is improved tenderness. Unlike the toughening process, tenderization is highly non-uniform. This variation in tenderization produces inconsistencies in meat tenderness (Koochmariaie, 1996).

Tenderness Standards

Sensory Analysis

One method of evaluating beef tenderness is by human sensory analysis. Beef samples are cooked, then analyzed by a trained sensory panel. This method is accurate, but destructive, time-consuming, and expensive.

Warner-Bratzler Shear Force

Currently, meat scientists use the Warner-Bratzler shear (WBS) force as an objective measure of tenderness. The process involves preparing a 2.54-cm thick steak, and cooking to an internal temperature of 70 °C. The steak is cooled, and 1.27-cm cores are removed from the ribeye, parallel to the muscle fibers (AMSA, 1995). The cores are then sheared perpendicular to the muscle fibers on an Instron Universal Testing Machine (Instron Corp, Canton, MA). The force required to shear the core is recorded and taken

as measure of meat tenderness. Higher shear-force values indicate tougher meat.

Generally, beef steaks are aged for 14-21 days at 1°C prior to cooking to match industry practice. WBS has been used by meat scientists for several decades. The method is accurate, but lacks repeatability due to two factors:

- Even though cores must be removed randomly from the cooked steaks, general guidelines suggest that the user avoid spots with thick marbling or connective tissue. Manual error is introduced.
- Cores must be removed at an angle parallel to muscle fiber. Subjective human judgment again intervenes.

Slice Shear Force

Researchers at the United States Meat Animal Research Center in Clay Center, NB have developed a slice shear force (SSF) system to measure beef tenderness (Shackelford et al., 1999a). The system is based on the WBS procedure described above, but utilizes belt-driven, automated equipment to increase speed and efficiency. For beef ribeye, the muscle fibers are at approximately 45° to the surface. SSF procedures require that the sample to be sheared be removed at a 45° angle, parallel to the muscle fibers. Instead of removing several cores, SSF removes one large sample and is then sheared. Koochmaraie (1999) and Wheeler et al. (2001) outline the following procedures to obtain the sample for shearing.

A first cut is made 1 cm from the lateral end of the cooked steak. Then, using box A (Fig. 2.6), the second cut is made 5 cm from the first cut. Utilizing slice box B, the slice shear sample is removed at an angle of 45° using a knife with two parallel blades spaced 1 cm apart. The cut is approximately parallel to muscle fiber orientation.

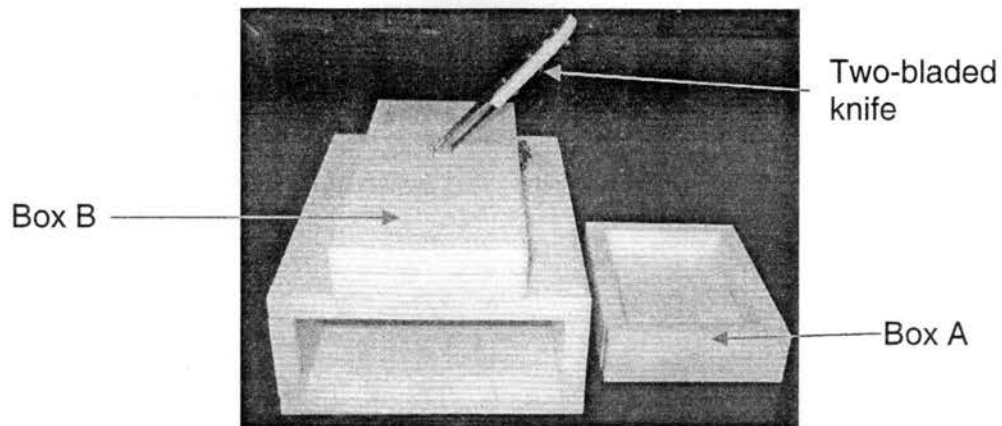


Figure 2.6. Sample slice acquisition equipment.

This procedure generates a cooked meat sample measuring 5 cm in length by 1 cm in thickness and 2.5 cm in width (Fig. 2.7). Sample location was selected to minimize connective tissue within the sample slice.

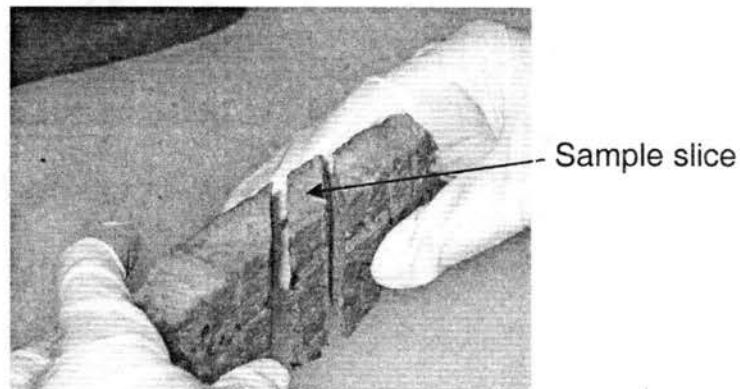


Figure 2.7. Sample slice from ribeye.

Slice shear force is measured using a flat, blunt-end blade attached to an Instron Universal Testing Machine. Force required to shear the muscle fibers of the slice is recorded as a slice shear force. Higher slice shear force values indicated “tougher” beef.

Shear-force values are obtained at three days postmortem and regressed using pre-determined regression equations to predict shear force at 14 days postmortem. The entire process (steak removal, cooking, shear-force determination, and classification) can be

completed in 10-15 minutes (Shackelford et al., 1997). The beef industry has not adopted this technique, because it is destructive to sample. However, meat scientists are beginning to prefer SSF values as a tenderness reference over WBS values, because of higher repeatability (NCBA, 2002).

Methods for Predicting Tenderness

Tenderness Probes

Various minimally invasive probe-type technologies have been developed. Swatland et al. (1998) developed a connective tissue probe that measures beef tenderness on the basis of ultraviolet fluorescence of connective tissue and the electro-mechanical resistance to penetration of the probe needle. For strip-loin steaks aged 21 days, they reported R^2 values for prediction of tenderness and flavor of 0.18 and 0.17, respectively. Jeremiah and Phillips (2000) evaluated the performance of a tenderness probe developed by the Meat Industry Research Institute of New Zealand and reported an R^2 value for prediction of overall tenderness of strip-loin steaks of 0.44.

George et al. (1997) evaluated an electromechanical penetrometer by Tendertec (Tendertec International, Bemboka, Australia). Tendertec predicted sensory tenderness scores with an R^2 value of 0.02. The authors concluded that this instrument was not useful for predicting beef tenderness.

Belk et al. (2001) evaluated Tendertec in four experiments. In the first two experiments, Tendertec did not correlate with WBS tenderness values. In the other two experiments, the correlations were significantly different from zero, but were weak. They concluded that Tendertec was unable to predict tenderness in young carcasses, and may be of limited use in older mature cattle. Smith (1999) reviewed different probe

technologies and reported that none of these probes predicted beef tenderness with sufficient accuracy.

Colorimeter

Wulf et al. (1997) investigated the correlation of colorimetric readings (L^* , brightness; a^* , red-green; b^* , yellow-blue) with beef tenderness, and reported that b^* values showed the highest correlation with shear force ($R^2 = 0.14$). Wulf and Wise (1999) reported that L^* , a^* , and b^* readings correlated with lean maturity scores by -0.67, -0.30, and -0.40, respectively.

Wulf and Page (2000) developed a quality grade augmentation system that supplemented current USDA Quality Grade Standards with $L^*a^*b^*$ colorimeter readings, pH, and hump height. In their study, they found that the darkest 20-25% of the beef carcasses were less palatable than others. Muscle color readings, when combined with marbling score, hump height, and muscle pH explained 36-46% of variation in beef palatability.

BeefCam[®]

Based on the color-tenderness relationship described above, Colorado State University in collaboration with Hunter Associates Laboratory, Inc. (Reston, VA) developed BeefCam[®], a computer vision-based system to predict beef tenderness. Wyle et al. (1999) evaluated the BeefCam[®] system in two trials. The R^2 value of prediction of tenderness for trials 1 and 2 were 0.19 and 0.07, respectively.

Vote et al. (2003) evaluated BeefCam[®] in four experiments. The R^2 values for predicting WBS values were 0.17, 0.30, 0.19, and 0.06 for the four tests. Combining the four experiments, 36% of tested carcasses were tough (WBS value > 4.5 kg). When

carcasses were ranked from tender to tough on the basis of BeefCam[®] predictions, only 3% of tough carcasses were correctly identified in the predicted top 10% of the toughest carcasses.

Wyle et al. (2003) evaluated BeefCam[®] using 769 carcasses. The regression model using variables measured by BeefCam[®] predicted WBS values with an R^2 value of 0.18. When quality grade data were included in the model, the R^2 value improved to 0.21. Of 769 carcasses tested, 13.8% of carcasses were tough. However, only 51.9% of carcasses were certified as tender, which indicates that large numbers of tender carcasses were misclassified as tough. They concluded that further development of the system was necessary for commercial application.

Slice Shear Force

In two experiments, Shackelford et al. (1997) evaluated 1-d or 2-d postmortem WBS values as a predictor of 14-d postmortem WBS values. The first experiment (n=400) was conducted in the laboratory, and the second experiment (n=554) was conducted under commercial processing conditions. They classified steaks with 1-d or 2-d WBS values into three tenderness categories (tender, medium, and tough). Steaks with 14-d WBS values were classified into two tenderness categories. All carcasses in the “tender” category, 1-d postmortem, were still in the “tender” category at 14-d postmortem. Most of the carcasses (81% and 85% in exp. 1 and 2, respectively) in the “medium” category, 1-d postmortem, were in the “tender” category at 14-d postmortem. Most of the carcasses (74% and 67% in exp. 1 and 2, respectively) in the “tough” category, day 1, were still in the “tough” category on day 14. Day-2 WBS values predicted 14-d WBS values with an R^2 value of 0.46.

Following the above study, Shackelford et al. (1999a) developed a system called “slice shear force” (SSF) for measuring beef tenderness under commercial processing conditions. Their SSF procedures were explained earlier. They reported an SSF repeatability of 0.89, exceeding the WBS repeatability of 0.85 reported earlier by Wheeler et al. (1998). Three-day postmortem SSF values accounted for 51% and 61% of variation in 14-d WBS values and sensory panel tenderness scores, respectively.

Shackelford et al. (1999b) compared WBS and SSF for predicting trained sensory panel tenderness ratings. SSF predicted sensory tenderness with an R^2 value of 0.67, whereas the corresponding R^2 value for WBS was 0.59. They reported that the SSF procedure is highly repeatable (0.91).

In a study supported by the National Cattlemen's Beef Association (Wheeler et al., 2002), the SSF, the colorimeter, and the BeefCam[®] were tested. The authors drew following conclusions:

- BeefCam[®] did not perform well in this study.
- The colorimeter performed inconsistently, giving unreliable results with USDA Select carcasses.
- Indirect, non-invasive methods to predict meat tenderness are not currently sufficiently accurate to warrant adoption by the industry.
- The direct measure of tenderness provided by SSF resulted in relatively accurate identification of “tender” beef carcasses. However, it appears that the industry does not favor this sample-destructive technology.

Results of the Wheeler et al. (2002) study are shown in Figure 2.8. At the 100% tenderness certification level, all carcasses were certified as tender, and therefore no

sorting was performed. Just over 9% of the tested carcasses were tough. The carcasses were ordered on the basis of predicted tenderness scores. At 90% tenderness certification, 10% of tough carcasses were grouped as “not certified tender.” The remaining 90% of carcasses were grouped as “certified tender.” Error was defined as percentage of tough carcasses in the “certified tender” group. If the instruments had performed optimally, all tough carcasses (about 9%) would be placed in the top 10%, and there would be no error in 90-10% tenderness certification levels. None of the systems met these criteria (Fig. 2.8).

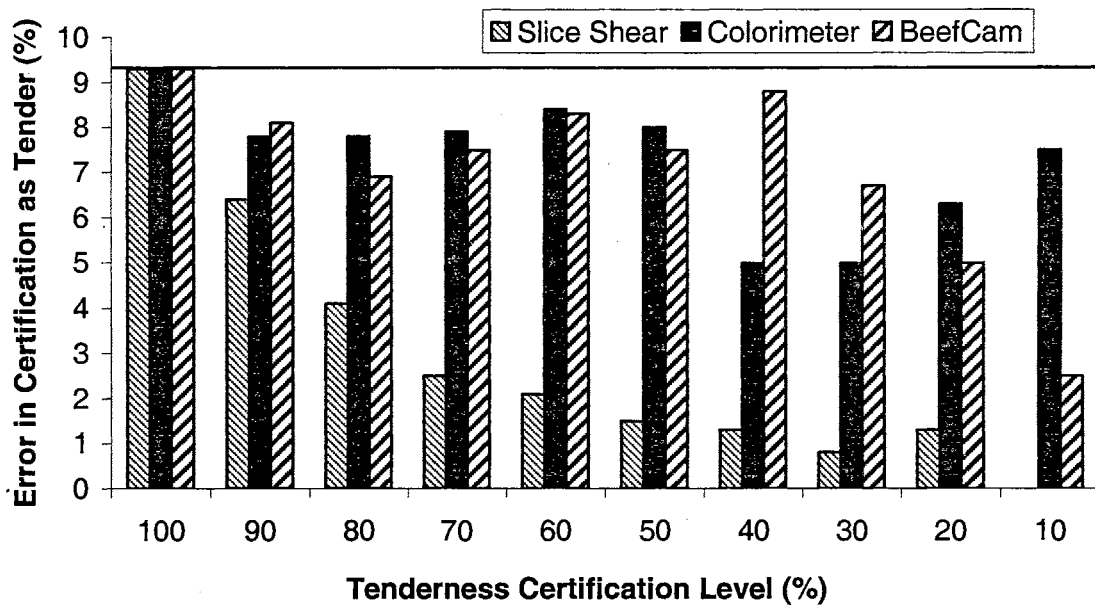


Figure 2.8. Comparison of SSF, Colorimeter, and BeefCam® (Wheeler et al., 2002).

Image Texture

Li et al. (2001) used wavelet-based textural features to classify beef carcasses into tender and tough groups. Fifty-nine samples were imaged with a spatial resolution of 480 x 512. Each image was divided into 56 subimages, with 64 x 64 resolution. From 8 extremely tender samples, 45 subimages were randomly selected. Another 45 subimages

were randomly selected from 9 extremely tough samples. These 90 subimages were the only images used in the study (Li et al., 1999a). It is not clear why the authors did not use more subimages, when one tough sample can produce 56 subimages. Wavelet textural features (n=35) were extracted from each subimage. Fisher's discriminant model was developed to classify tender and tough samples and was evaluated by cross-validation (leave-one-out technique). They achieved 83.3% correct classification (Li et al., 2001; Li et al., 1999a). Only 17 samples with extreme tenderness scores were used in this study. Had the other samples with intermediate tenderness scores been used, the correct classification rate would obviously have been reduced.

Li et al. (1999b) showed the utility of image textural features for predicting beef tenderness from 90 imaged samples. Color and marbling features predicted sensory tenderness scores with an R^2 value of 0.30. When textural features were included, the R^2 value increased to 0.72.

Near-infrared Spectroscopy

Light reflected in the visible region of the spectrum gives an objective measurement of the color of food objects, whereas light reflected in the near-infrared (NIR) region contains information about food physical and chemical properties. Contradictory results have been reported in the literature regarding the use of NIR spectroscopy for predicting beef tenderness. Some studies have been conducted using laboratory spectrometers with special requirements for sample preparation. Others have used fiber-optic reflectance probes that can be more easily applied in a processing plant environment. In most studies, the NIR spectral scan and reference shear-force tenderness

values were acquired at the same point in time. A few studies have used scans acquired at early postmortem to predict tenderness after aging. It is important to recognize these differences in order to understand the inconsistent results reported in the literature.

Pioneering Work

Use of NIR spectroscopy for predicting beef tenderness was pioneered by Mitsumoto et al. (1991). A limited number of samples was used in the study (11 steers with 6 muscles/steer). They collected spectral reflectance data (1100-2500 nm) on 3-d postmortem samples to predict 3-d WBS values. Samples were prepared and placed in sample holders for scanning. A multiple linear regression equation with four selected wavelengths predicted 3-d WBS values with R^2 values of 0.68. This study showed promise and aroused broad interest in using NIR for predicting beef tenderness.

Matforsk Norwegian Research Group

A series of studies were conducted at the Matforsk Norwegian Food Research Institute. Hildrum et al. (1994) conducted a feasibility study with 10 carcasses (3 cows, 7 bulls) at 3 aging periods (1, 8, and 14 days). They acquired scans in 1100-2500 nm range under controlled conditions requiring special sample preparation. Thirty spectral scans were used to build a two-factor principal component regression model with multiplicative scatter correction (MSC) data preprocessing. The model predicted WBS values with an R^2 value of 0.62. Note that 1-d spectra predicted 1-d WBS values, while 14-d spectra predicted 14-d WBS values.

In continuation of the work, Hildrum et al. (1995) used 40 carcasses (both cows and bulls) at 3 aging periods (1, 7, and 14 days). Thirty carcasses were chilled slowly, whereas 10 carcasses were subjected to electric stimulation and rapid chilling. Spectra

were collected from fresh samples and frozen/thawed samples. It is interesting to note that the NIR spectra collected from frozen/thawed samples predicted WBS values more accurately ($R^2 = 0.50$) than those collected fresh ($R^2 = 0.34$). The authors speculated that the freezing step increased the difference in structure between tender and tough samples. Note that the model predicted current WBS values rather than forecasting future WBS values.

Rødbotten et al. (2000) studied the feasibility of using early (4- and 26-hour) postmortem NIR spectroscopy to predict 2-d and 7-d WBS values from 127 carcasses. Spectra recorded at 4-h and 26-h postmortem exhibited some differences in 1100-1900 nm range. The rigor mortis process starts at about 4-h postmortem and ends before 26 hours. The authors hypothesized that the difference between the two spectra collected at the beginning and end of rigor mortis is related to the aging potential of the carcasses. However, they found that spectra differences did not contain essential information to predict final tenderness of aged samples. Partial least squares (PLS) models with MSC preprocessing of 4-h NIR spectra predicted 2-d and 7-d WBS readings with R^2 values of 0.25 and 0.27, respectively. Models based on 26-h spectra predicted 2-d and 7-d WBS readings with R^2 values of 0.36 and 0.37, respectively. They concluded that their study did not support the hypothesis that early postmortem NIR spectroscopy can be used as a predictor of final tenderness.

Rødbotten et al. (2001) used a spectrometer (950-1700 nm) equipped with a diode array detector to predict tenderness of 12 carcasses. Traditional NIR spectrometers use grating and require scanning to acquire complete spectra. Diode array-based spectrometers acquire spectra without scanning, and thus acquisition is more rapid. The

authors were able to acquire 150 spectra to cover the entire surface of a sample on both sides in 7 s. Four samples were collected from each carcass. NIR spectra and WBS readings were recorded at 2, 9, and 21 days postmortem. Separate models for different aging periods predicted current tenderness with R^2 value of 0.36-0.69. With regard to forecasting tenderness, a 2-d spectrum model predicted 9-d WBS values with moderate accuracy ($R^2 = 0.52$). Less accuracy was reported for forecasting 21-d WBS values ($R^2 = 0.27$). Special sample preparation was not required for this spectrometer.

Teagasc Irish Research Group

Byrne et al. (1998) used a scanning spectrometer (750-1098 nm) equipped with a fiber optic probe to predict beef tenderness of 70 heifer carcasses. Due to lamp replacement during the course of this study, two separate models were developed using 20 and 50 samples, respectively. Results of the second set are described here. A 10-factor PLS model based on 1-d spectra predicted 14-d WBS values with $R^2 = 0.68$. Models based on 2-, 7-, and 14-d spectra predicted less accurately ($R^2 = 0.20 - 0.45$). This result is intriguing, because 1-d spectra were able to predict 14-d WBS values more accurately than 14-d spectra. This result contradicts expectations, because a 1-d spectrum is forecasting tenderness that involves the variability of aging, whereas a 14-d spectrum predicts current tenderness. Even the prediction accuracies of models based on 1-d and 2-d spectra were considerably different. NIR readings and WBS values were collected on different steaks from the same carcasses. The authors suggested that the collecting NIR readings and WBS values from the same meat sample would improve the accuracy.

Continuing the above work, Venel et al. (2001) acquired NIR spectra (750-1098 nm) on *longissimus dorsi* (l.d.) and *semimembranosus* muscles from 75 animals.

Following the recommendation made by Byrne et al. (1998), the site of spectral measurement was marked with a scalpel, and WBS values were measured at the same site. Spectra and WBS values were collected at 14-d post mortem. NIR spectroscopy was unable to predict tenderness for *semimembranosus* muscle. For l.d. muscle, R^2 values for tenderness prediction ranged from 0.27-0.33. When separate models were developed for each segregation of samples based on sex, grade, and pH, R^2 values of prediction ranged from 0.15-0.54.

Other Work

Park et al. (1998) collected NIR spectra (1100-2498 nm) and WBS values on frozen and thawed samples at either 7-d or 14-d post mortem on 119 samples. Special sample preparation was required. They developed a 6-factor PLS model to predict tenderness with R^2 values of 0.67 and 0.63 for calibration (n=80) and validation (n=39) sets, respectively. Note that 7-d spectra predicted 7-d WBS values and 14-d spectra predicted 14-d WBS values. Hildrum et al. (1995) reported that the spectra collected on the samples that were frozen/thawed predicted WBS values more accurately than those collected on fresh samples. However, samples will not be frozen and thawed in industry application.

Liu et al. (2003) collected NIR spectra (400-2498 nm) and WBS values from 24 carcasses at 2, 4, 8, 14, and 21 days postmortem. An 8-factor PLS model predicted WBS values with $R^2 = 0.49$. Note that the model did not forecast aged tenderness. When the data were segregated into different aging periods, prediction accuracy increased for 4-d and 21-d aging periods, while accuracy decreased for other aging periods.

Leroy et al. (2003) used a Fourier Transform NIR spectrometer to record spectra (833-2500 nm) from 189 samples using transmission and reflectance modes at 2-d and 8-d postmortem. Using the reflectance mode, the 2-d spectra model predicted 2-d and 8-d WBS readings with R^2 values of 0.25 and 0.19, respectively. With the transmission mode, a 2-d spectra model predicted 2-d and 8-d WBS readings with R^2 values of 0.41 and 0.15, respectively.

To summarize, several studies reported moderate to promising results in predicting current-status tenderness (Mitsumoto et al., 1991; Hildrum et al., 1994, 1995; Rødbotten et al., 2001; Park et al., 1998; Liu et al., 2003). Only one study (Byrne et al., 1998) reported success in forecasting tenderness. However, another study from the same research group (Venel et al., 2001) reported failure in predicting current tenderness. Three studies reported failure in forecasting tenderness (Rødbotten et al., 2000, 2001; Leroy et al., 2003).

X-ray Absorption

Dual-energy X-ray absorption (DEXA) has been used successfully for body composition analysis in humans and animals. Principles of DEXA have been demonstrated by Buzzell and Pintauro (2004). Mitchell et al. (1997) investigated dual-energy X-ray absorption for composition analysis of three-rib sections of beef using a commercial densitometer designed for medical imaging. They took scans at energy levels of 38 and 70 keV. Reported R-values (ratio of attenuation coefficient at the two energy levels) were 1.2 for fat and 1.4 for 100% lean (Mitchell et al., 2003). Calibration details were not given. DEXA measurements predicted the compositions of ribs with correlation coefficients ranging from 0.85 to 0.94. Continuing the work, Mitchell et al.

(1998a) used DEXA to study composition of pork carcasses. Mitchell et al. (1998b, 1998c, 2001b) used DEXA for body composition analysis and to measure bone mineral content and density in live pigs.

Brienne et al. (2001) used DEXA for assessment of meat fat content. They used a commercial densitometer that employs 24 sodium iodide detectors, to obtain images at 44 and 70 keV at 0.2 mA. They positioned a series of 12 Plexiglass plates under the X-ray source tube to attenuate the X-ray photons and avoid detector saturation. Then, they estimated incident intensity, I_0 , by accounting for the absorption due to Plexiglass.

DEXA predicted various chemical components with R^2 values ranging from 0.70 to 0.97.

A patent application has been filed in New Zealand covering evaluation of meat properties (including tenderness) using X-ray absorption. The patent disclosure indicates that the invention does not use an objective test of meat tenderness for comparison with X-ray measurements. Their “estimated tenderness” is based on unidentified visual features of the meat (Murray, 2001).

References

- Allen, P., and N. Finnerty. 2000. Objective beef classification: A report of a trial of three VIA classification systems. Dublin, Ireland: The National Food Centre, Teagasc. Available at: <http://www.teagasc.ie/publications/beefgrading/beefgrading.pdf>. Accessed on 16 January 2004.
- AMSA. 1995. *Research guidelines for cookery, sensory evaluation, and instrumental tenderness measurements of fresh meat*. Chicago, IL.: National Live Stock & Meat Board, American Meat Science Association.

- Antequera, T., E. Muriel, P. G. Rodríguez, E. Cernadas, and J. Ruiz. 2003. Magnetic resonance imaging as a predictive tool for sensory characteristics and intramuscular fat content of dry-cured loin. *Journal of the Science of Food and Agriculture* 83(4): 268-274.
- Belk, K. E., M. H. George, J. D. Tatum, G. G. Hilton, R. K. Miller, M. Koohmaraie, J. O. Reagan, and G. C. Smith. 2001. Evaluation of the Tendertec beef grading instrument to predict the tenderness of steaks from beef carcasses. *Journal of Animal Science* 79(3): 688-697.
- Belk, K. E., J. A. Scanga, J. D. Tatum, J. W. Wise, and G. C. Smith. 1998. Simulated instrument augmentation of USDA yield grade application to beef carcasses. *Journal of Animal Science* 76(2): 522-527.
- Benn, A., D. Barrett-Lennard, and P. J. Hay. 1998. Image analysis for meat. U.S. Patent No. 5793879.
- Bertram, H. C., A. K. Whittaker, W. R. Shorthose, H. J. Andersen, and A. H. Karlsson. 2004. Water characteristics in cooked beef as influenced by ageing and high-pressure treatment--an NMR micro imaging study. *Meat Science* 66(2): 301-306.
- Boleman, S. J., S. L. Boleman, R. K. Miller, J. F. Taylor, H. R. Cross, T. L. Wheeler, M. Koohmaraie, S. D. Shackelford, M. F. Miller, R. L. West, D. D. Johnson, and J. W. Savell. 1997. Consumer evaluation of beef of known categories of tenderness. *Journal of Animal Science* 75(6): 1521-1524.

- Boleman, S. L., S. J. Boleman, W. W. Morgan, D. S. Hale, D. B. Griffin, J. W. Savell, R. P. Ames, M. T. Smith, J. D. Tatum, T. G. Field, G. C. Smith, B. A. Gardner, J. B. Morgan, S. L. Northcutt, H. G. Dolezal, D. R. Gill, and F. K. Ray. 1998. National Beef Quality Audit-1995: survey of producer-related defects and carcass quality and quantity attributes. *Journal of Animal Science* 76(1): 96-103.
- Bonny, J.-M., W. Laurent, and J.-P. Renou. 2000a. Detection of susceptibility effects using simultaneous T2* and magnetic field mapping. *Magnetic Resonance Imaging* 18(9): 1125-1128.
- Bonny, J.-M., W. Laurent, R. Labas, R. Taylor, P. Berge, and J.-P. Renou. 2000b. Magnetic resonance imaging of connective tissue: a non-destructive method for characterising muscle structure. *Journal of the Science of Food and Agriculture* 81(3): 337 - 341.
- Borggaard, C., N. T. Madsen, and H. H. Thodberg. 1996. In-line image analysis in the slaughter industry, illustrated by Beef Carcass Classification. *Meat Science* 43(Supplement 1): 151-163.
- Brienne, J. P., C. Denoyelle, H. Baussart, and J. D. Daudin. 2001. Assessment of meat fat content using dual energy X-ray absorption. *Meat Science* 57(3): 235-244.
- Buzzell, P., and S. Pintauro. 2004. Dual energy X-ray absorptiometry. Burlington, Vermont: Department of Nutrition and Food Sciences, University of Vermont. Available at: <http://nutrition.uvm.edu/bodycomp/dexa/>. Accessed on 10 February 2004.

- Byrne, C. E., G. Downey, D. J. Troy, and D. J. Buckley. 1998. Non-destructive prediction of selected quality attributes of beef by near-infrared reflectance spectroscopy between 750 and 1098 nm. *Meat Science* 49(4): 399-409.
- Cannell, R. C., J. D. Tatum, K. E. Belk, J. W. Wise, R. P. Clayton, and G. C. Smith. 1999. Dual-component video image analysis system (VIASCAN) as a predictor of beef carcass red meat yield percentage and for augmenting application of USDA yield grades. *Journal of Animal Science* 77(11): 2942-2950.
- Cross, H. R., and K. E. Belk. 1994. Objective measurements of carcass and meat quality. *Meat Science* 36(1-2): 191-202.
- Cross, H. R., D. A. Gilliland, P. R. Durland, and S. C. Seideman. 1983. Beef carcass evaluation by use of a video image analysis system. *Journal of Animal Science* 57: 908-917.
- Cross, H. R., and A. D. Whittaker. 1992. The role of instrument grading in a beef value-based marketing system. *Journal of Animal Science* 70: 984-989.
- George, M. H., J. D. Tatum, H. G. Dolezal, J. B. Morgan, J. W. Wise, C. R. Calkins, T. Gordon, J. O. Reagan, and G. C. Smith. 1997. Comparison of USDA quality grade with tendertec for the assessment of beef palatability. *Journal of Animal Science* 75(6): 1538-1546.
- Gerrard, D. E., X. Gao, and J. Tan. 1996. Beef marbling and color score determination by image processing. *Journal of Food Science* 61(1): 145-148.
- Goering, K. 1999. Instrument grading: A different sort of measurement. *National Cattlemen*, 14(1): 54-56.

- Hammer, B. E. 1998. Industrial applications of nuclear magnetic resonance. *Sensor Review* 18(4): 245-251.
- Hatem, I., J. Tan, and D. E. Gerrard. 2003. Determination of animal skeletal maturity by image processing. *Meat Science* 65(3): 999-1004.
- Hatem, I., and J. Tan. 2003. Cartilage and bone segmentation in vertebra images. *Transactions of the ASAE* 46(5): 1429-1434.
- Hildrum, K. I., T. Isaksson, T. Neas, B. N. Nilsen, M. Rødbotten, and P. Lea. 1995. Near infrared reflectance spectroscopy in the prediction of sensory properties of beef. *Journal of Near Infrared Spectroscopy* 3: 81-87.
- Hildrum, K. I., B. N. Nilsen, M. Mielnik, and T. Naes. 1994. Prediction of sensory characteristics of beef by near-infrared spectroscopy. *Meat Science* 38(1): 67-80.
- Ishmael, W. 2000. Value Vision. Available at: http://beef-mag.com/ar/beef_value_vision/index.htm. Accessed on 6 February 2004.
- Jeremiah, L. E. 1996. The influence of subcutaneous fat thickness and marbling on beef : Palatability and consumer acceptability. *Food Research International* 29(5-6): 513-520.
- Jeremiah, L. E., and D. M. Phillips. 2000. Evaluation of a probe for predicting beef tenderness. *Meat Science* 55(4): 493-502.
- Koohmaraie, M. 1995. The biological basis of meat tenderness and potential genetic approaches for its control and prediction. Clay Center, NE.: USDA-ARS, Roman L. Hruska U.S. Meat Animal Research Center. Available at: http://meats.marc.usda.gov/MRU_WWW/ICMST95/ICMST95.html. Accessed on 9 February 2004.

- Koohmaraie, M. 1996. Biochemical factors regulating the toughening and tenderization processes of meat. Clay Center, NE.: USDA-ARS, Roman L. Hruska U.S. Meat Animal Research Center,. Available at:
http://meats.marc.usda.gov/MRU_WWW/ICMST96/ICMST96.html. Accessed on 9 February 2004.
- Koohmaraie, M. 1999. MARC beef classification system: objective evaluation of beef tenderness and cutability. Clay Center, NE.: Roman L. Hruska U. S. Meat Animal Research Center, Agricultural Research Service, United States Department of Agriculture. Available at: http://meats.marc.usda.gov/MRU_WWW/Bulletin.pdf. Accessed on 17 February 2004.
- Lenhert, D. H., and D. A. Gilliland. 1985. The design and testing of an automated beef grader. ASAE Paper No. 85-3035. St. Joseph, Mich.: ASAE.
- Leroy, B., S. Lambotte, O. Dotreppe, H. Lecocq, L. Istasse, and A. Clinquart. 2003. Prediction of technological and organoleptic properties of beef Longissimus thoracis from near-infrared reflectance and transmission spectra. *Meat Science* 66(1): 45-54.
- Li, J., J. Tan, and P. Shatadal. 1999a. Discrimination of beef images by textural features. ASAE Paper No. 99-3158. St. Joseph, Mich.: ASAE.
- Li, J., J. Tan, F. A. Martz, and H. Heymann. 1999b. Image texture features as indicators of beef tenderness. *Meat Science* 53(1): 17-22.
- Li, J., J. Tan, and P. Shatadal. 2001. Classification of tough and tender beef by image texture analysis. *Meat Science* 57(4): 341-346.

- Liu, Y., B. G. Lyon, W. R. Windham, C. E. Realini, T. D. D. Pringle, and S. Duckett. 2003. Prediction of color, texture, and sensory characteristics of beef steaks by visible and near infrared reflectance spectroscopy. A feasibility study*. *Meat Science* 65(3): 1107-1115.
- Lu, W., and J. Tan. 2004. Analysis of image-based measurements and USDA characteristics as predictors of beef lean yield. *Meat Science* 66(2): 483-491.
- Lusk, J. L., J.A. Fox, T.C. Schroeder, J. Mintert, and M. Koohmaraie. 2001. In-store valuation of steak tenderness. *American Journal of Agricultural Economics* 83(3): 539-550.
- Madsen, N. T., and H. H. Thodberg. 1999. Objective beef classification and quality measurements. In *European experts colloquium on grading and classification*. Kulmbach, Germany: European Commission Fair, MACA Meeting.
- McKenna, D. R., D. L. Roebert, P. K. Bates, T. B. Schmidt, D. S. Hale, D. B. Griffin, J. W. Savell, J. C. Brooks, J. B. Morgan, T. H. Montgomery, K. E. Belk, and G. C. Smith. 2002. National Beef Quality Audit-2000: survey of targeted cattle and carcass characteristics related to quality, quantity, and value of fed steers and heifers. *Journal of Animal Science* 80(5): 1212-1222.
- Mitchell, A. D., A. M. Scholz, V. G. Pursel, and C. M. Evock-Clover. 1998a. Composition analysis of pork carcasses by dual-energy x-ray absorptiometry. *Journal of Animal Science* 76(8): 2104-2114.
- Mitchell, A. D., A. M. Scholz, and J. M. Conway. 1998b. Body composition analysis of small pigs by dual-energy x-ray absorptiometry. *Journal of Animal Science* 76(9): 2392-2398.

- Mitchell, A. D., A. M. Scholz, and J. M. Conway. 1998c. Body composition analysis of pigs from 5 to 97 kg by dual energy X-ray absorptiometry. *Applied Radiation Isotopes*, 49: 521-523.
- Mitchell, A. D., A. M. Scholz, P. C. Wange, and H. Song. 2001a. Body composition analysis of the pig by magnetic resonance imaging. *Journal of Animal Science* 79(7): 1800-1813.
- Mitchell, A. D., A. M. Scholz, and V. G. Pursel. 2001b. Total body and regional measurements of bone mineral content and bone mineral density in pigs by dual energy X-ray absorptiometry. *Journal of Animal Science* 79(10): 2594-2604.
- Mitchell, A. D., A. M. Scholz, and V. G. Pursel. 2003. Prediction of pork carcass composition based on cross-sectional region analysis of dual energy X-ray absorptiometry (DXA) scans. *Meat Science* 63(2): 265-271.
- Mitchell, A. D., M. B. Solomon, and T. S. Rumsey. 1997. Composition analysis of beef rib sections by dual-energy X-ray absorptiometry. *Meat Science* 47(1-2): 115-124.
- Mitsumoto, M., S. Maeda, T. Mitsuhashi, and S. Ozawa. 1991. Near-infrared spectroscopy determination of physical and chemical characteristics in beef cuts. *Journal of Food Science* 59(6): 1493-1496.
- Morgan, J. B., J. W. Savell, D. S. Hale, R. K. Miller, D. B. Griffin, H. R. Cross, and S. D. Shackelford. 1991. National beef tenderness survey. *Journal of Animal Science* 69(8): 3274-3283.
- Murray, B. C. 2001. A method for the non-invasive measurement of properties of meat. International Publication No. WO 01/96844.

- NCBA. 2002. Beef Update. Meeting Summary. National Beef Instrument Assessment Plan II: Focus on Tenderness. Centennial, CO: National Cattlemen's Beef Association. Available at:
<http://www.beefboard.org/documents/beef%20update.pdf>. Accessed on 9 February 2004.
- Nielson, T., N. T. Madsen, and H. H. Thodberg. 1997. Technology transfer in computer vision - Example: beef carcass classification centre BCC-2. In *Technology Transfer in Computer Vision*. Lappeeranta, Finland: ECVnet.
- Park, B., Y. R. Chen, W. R. Hruschka, S. D. Shackelford, and M. Koohmaraie. 1998. Near-infrared reflectance analysis for predicting beef longissimus tenderness. *Journal of Animal Science* 76(8): 2115-2120.
- Perkins, T., A. Meadows, and B. Hays. 2003. Study Guide for the Ultrasonic evaluation of beef cattle for carcass merit. Ames, IA.: Ultrasound Guidelines Council, Beef Cattle Ultrasound Technician Annual Proficiency Testing and Certification, Iowa State University. Available at:
<http://www.aptcbeef.org/UGC%20STUDY%20GUIDE.pdf>. Accessed on 11 February 2004.
- Petersen, F., S. Klastrup, N. T. Madsen, and S. E. Sorensen. 1989. Beef Classification Center. In *Proceedings of the International Congress of Meat Science and Technology*, 35: 49-52.
- Renou, J. P., L. Foucat, and J. M. Bonny. 2003. Magnetic resonance imaging studies of water interactions in meat. *Food Chemistry* 82(1): 35-39.

- Rødbotten, M., B. N. Nilsen, and K. I. Hildrum. 2000. Prediction of beef quality attributes from early post mortem near infrared reflectance spectra. *Food Chemistry* 69(4): 427-436.
- Rødbotten, R., B.-H. Mevik, and K. I. Hildrum. 2001. Prediction and classification of tenderness in beef from non-invasive diode array detected NIR spectra. *Journal of Near Infrared Spectroscopy* 9: 199-210.
- Shackelford, S. D., T. L. Wheeler, and M. Koohmaraie. 1997. Tenderness classification of beef: I. Evaluation of beef longissimus shear force at 1 or 2 days postmortem as a predictor of aged beef tenderness. *Journal of Animal Science* 75(9): 2417-2422.
- Shackelford, S. D., T. L. Wheeler, and M. Koohmaraie. 1998. Coupling of image analysis and tenderness classification to simultaneously evaluate carcass cutability, longissimus area, subprimal cut weights, and tenderness of beef. *Journal of Animal Science* 76(10): 2631-2640.
- Shackelford, S. D., T. L. Wheeler, and M. Koohmaraie. 1999a. Tenderness classification of beef: II. Design and analysis of a system to measure beef longissimus shear force under commercial processing conditions. *Journal of Animal Science* 77(6): 1474-1481.
- Shackelford, S. D., T. L. Wheeler, and M. Koohmaraie. 1999b. Evaluation of slice shear force as an objective method of assessing beef longissimus tenderness. *Journal of Animal Science* 77(10): 2693-2699.

- Shackelford, S. D., T. L. Wheeler, and M. Koohmaraie. 2003. On-line prediction of yield grade, longissimus muscle area, preliminary yield grade, adjusted preliminary yield grade, and marbling score using the MARC beef carcass image analysis system. *Journal of Animal Science* 81(1): 150-155.
- Shackelford, S. D., T. L. Wheeler, M. K. Meade, J. O. Reagan, B. L. Byrnes, and M. Koohmaraie. 2001. Consumer impressions of Tender Select beef. *Journal of Animal Science* 79(10): 2605-2614.
- Signe, K., P. Freddy, and S. S. Erik. 1989. Method and apparatus for the determination of quality properties of individual cattle carcasses. European Patent No. 0321981.
- Smith, G. C. 1999. New technologies for precision selection, management and marketing of beef. In *Beef Information Days*. Pullman, WA: Washington State University.
- Steiner, R., A. M. Wyle, D. J. Vote, K. E. Belk, J. A. Scanga, J. W. Wise, J. D. Tatum, and G. C. Smith. 2003a. Real-time augmentation of USDA yield grade application to beef carcasses using video image analysis. *Journal of Animal Science* 81(9): 2239-2246.
- Steiner, R., D. J. Vote, K. E. Belk, J. A. Scanga, J. W. Wise, J. D. Tatum, and G. C. Smith. 2003b. Accuracy and repeatability of beef carcass longissimus muscle area measurements. *Journal of Animal Science* 81(8): 1980-1988.
- Stroshine, R. 1998. *Physical Properties of Agricultural Materials and Food Products*. IN.: West Lafayette: Copy Cat.
- Swatland, H. J., J. C. Brooks, and M. F. Miller. 1998. Possibilities for predicting taste and tenderness of broiled beef steaks using an optical-electromechanical probe. *Meat Science* 50(1): 1-12.

- Tong, A. K., D. J. Robinson, and T. Liu. 1998. Method and apparatus for using image analysis to determine meat and carcass characteristics. Canadian Patent No. 02263763.
- USDA. 1997. United States standards for grades of carcass beef. Effective date January 31, 1997. Washington, DC.: Livestock and Seed Division of the Agricultural Marketing Service. Available at: <http://www.ams.usda.gov/lsg/stand/standards/beef-car.pdf>. Accessed on 9 February 2004.
- USDA. 2001. Procedures for approval and use of instrument augmentation systems for beef carcass ribeye measurement. Washington, DC.: Livestock and Seed Program, Agricultural Marketing Service.
- Venel, C., A. M. Mullen, G. Downey, and D. J. Troy. 2001. Prediction of tenderness and other quality attributes of beef by near infrared reflectance spectroscopy between 750 and 1100 nm; further studies. *Journal of Near Infrared Spectroscopy* 9: 185-198.
- Vote, D. J., K. E. Belk, J. D. Tatum, J. A. Scanga, and G. C. Smith. 2003. Online prediction of beef tenderness using a computer vision system equipped with a BeefCam module. *Journal of Animal Science* 81(2): 457-465.
- 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.

- Wheeler, T. L., S. D. Shackelford, and M. Koohmaraie. 2001. Shear force procedures for meat tenderness measurement. Clay Center, NE.: Roman L. Hruska U. S. Meat Animal Research Center, Agricultural Research Service, United States Department of Agriculture. Available at: http://meats.marc.usda.gov/MRU_WWW/Protocol/WBS.pdf. Accessed on 17 February 2004.
- Wheeler, T. L., S. D. Shackelford, and M. Koohmaraie. 1998. Cooking and palatability traits of beef longissimus steaks cooked with a belt grill or an open hearth electric broiler. *Journal of Animal Science* 76(11): 2805-2810.
- Wheeler, T. L., D. Vote, J. M. Leheska, S. D. Shackelford, K. E. Belk, D. M. Wulf, B. L. Gwartney, and M. Koohmaraie. 2002. The efficacy of three objective systems for identifying beef cuts that can be guaranteed tender. *Journal of Animal Science* 80(12): 3315-3327.
- Winger, R. C., and C. J. Hagyard. 1994. Chapter 4: Juiciness - its importance and some contributing factors. In *Quality attributes and their measurement in meat, poultry and fish products. Advances in meat research series*. Vol. 9. 94-124. A. M. Pearson and T. R. Dutson, eds. London, UK.: Blackie Academic & Professional.
- Wulf, D. M., S. F. O'connor, J. D. Tatum, and G. C. Smith. 1997. Using objective measures of muscle color to predict beef longissimus tenderness. *Journal of Animal Science* 75(3): 684-692.

- Wulf, D. M., and J. K. Page. 2000. Using measurements of muscle color, pH, and electrical impedance to augment the current USDA beef quality grading standards and improve the accuracy and precision of sorting carcasses into palatability groups. *Journal of Animal Science* 78(10): 2595-2607.
- Wulf, D. M., and J. W. Wise. 1999. Measuring muscle color on beef carcasses using the L*a*b* color space. *Journal of Animal Science* 77(9): 2418-2427.
- Wyle, A. M., R. C. Cannell, K. E. Belk, M. Goldberg, R. Riffle, and G. C. Smith. 1999. An evaluation of the prototype portable Hunterlab video imaging system (BeefCam) as a tool to predict tenderness of beef carcasses using objective measures of lean and fat color. 1999 Research Report. Fort Collins, CO.: Department of Animal Science, Colorado State University.
- Wyle, A. M., D. J. Vote, D. L. Roeber, R. C. Cannell, K. E. Belk, J. A. Scanga, M. Goldberg, J. D. Tatum, and G. C. Smith. 2003. Effectiveness of the SmartMV prototype BeefCam System to sort beef carcasses into expected palatability groups. *Journal of Animal Science* 81(2): 441-448.

CHAPTER III

COMPUTER VISION SEGMENTATION OF THE *LONGISSIMUS*

***DORSI* FOR BEEF QUALITY GRADING**

S. Jeyamkondan, N. Ray, G.A. Kranzler, and S. Acton

Abstract

A computer vision system was developed to support automation of beef quality grading. Images of beef steaks were acquired for algorithm development. Fat and lean were differentiated using a fuzzy c-means clustering algorithm. Segmentation of the longissimus dorsi (l.d.) muscle is required, because experts assign quality grades based primarily on visual appraisal of the l.d. A robust segmentation algorithm was developed using convex hull procedures. The l.d. was segmented from the steak using morphological operations of erosion and dilation. At the end of each iteration of erosion and dilation, a convex hull was fitted to the image, and compactness was measured. Iterations were continued to yield the most compact l.d. Classification error in segmentation was 1.97%. Average error pixel distance of segmentation by the computer vision system was 4.4 pixels.

Keywords. beef grading, adaptive segmentation, video image analysis, fuzzy c-means clustering, convex hull, computer vision.

Introduction

The beef industry is currently the largest food and fiber industry in the U.S., with \$55 billion of retail sales in 1999 (NCBA, 2000). To facilitate marketing, beef grading standards were developed by the United States Department of Agriculture (USDA) to classify carcasses into quality and yield grades. A trained employee of the USDA, working independently of both the producer and the packer, determines beef grades.

Graders assign quality grades from visual appraisal of the *longissimus dorsi* (l.d.) muscle from the carcass sectioned between the 12th and 13th rib (Fig. 3.1a). Quality grades are based on marbling levels (abundance and distribution of intramuscular fat) and physiological maturity of the carcass. Graders follow USDA quality grading standards (USDA, 1997).

Typical packing plant production rates allow only 9 to 18 seconds to assign grades (Belk et al., 1998). Although highly trained, these graders are subject to fatigue and emotional strain, which can affect the decision-making process. The subjective nature of the grading process produces inconsistencies in judgment. Resulting grading errors have negative economic impact on the producers as well as packers.

An objective “instrumented” carcass evaluation procedure is needed to ensure consistent quality and to facilitate value-based marketing. The Value-Based Marketing Task Force of the Cattlemen’s Beef Board and the National Beef Instrument Assessment Plan of the National Live Stock and Meat Board identified video image analysis (VIA) as the most promising technique for instrumenting carcass evaluation (Belk et al., 1996).

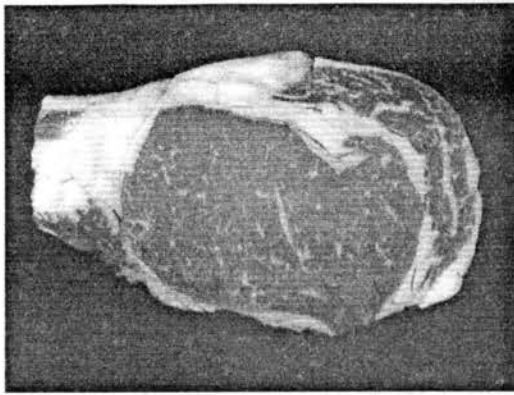
Computer vision/VIA has significant potential for automating the quality control process. Computer vision technology in the form of color video cameras, image

digitizers, and computer processing best matches the human eye for quality grading (Biju, 1998). Over the past 15 years, researchers have applied computer vision to the grading of various agricultural products such as tree seedlings (Kranzler and Rigney, 1989) and pork carcasses (Goldenberg and Lu, 1997). There is interest around the globe in applying computer vision to beef grading (Belk et al., 2001; Benn et al., 1998; Borggaard et al., 1996; Newman, 1995; Tong et al., 1998).

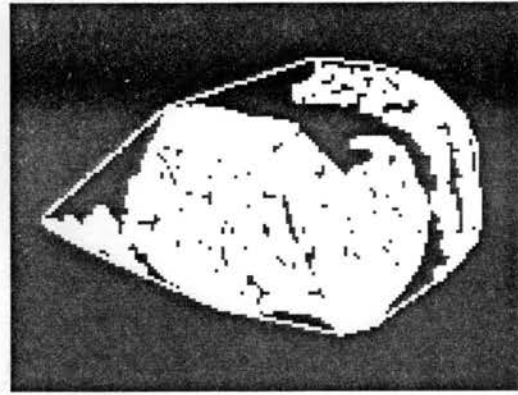
Segmentation of the l.d. muscle from the steak image (Fig. 3.1) is the primary step in developing a successful computer vision system for grading beef. Small errors in segmentation can lead to incorrect interpretation of intermuscular fat as marbling and result in considerable error in determination of marbling score and quality grade.

McDonald and Chen (1990) used the morphological operation, 'open,' to remove extraneous tissue from the l.d. muscle. Computation time was 30 min for the morphological operations. In the Canadian Vision System, a commercially available beef grading system, a heuristic approach was implemented to define the l.d. muscle boundary (Tong et al., 1998). Accuracy of segmentation was not reported.

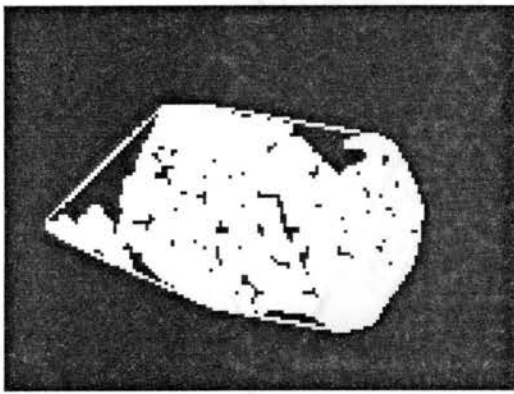
Gao et al. (1995) described two image segmentation techniques. They first employed rectangular prism boundary methods. Their second method was based on Mahalanobis distance in red-green-blue (RGB) color space. Average grade classification errors for five ham images were 16% and 15% for the first and second techniques, respectively. Lu and Tan (1998) reported that the segmentation of lean from fat can be improved by using non-linear transformations such as quadratic and hyperbolic tangent transforms on pixel values. However, the authors did not report removal of extraneous muscle tissue that was not part of the l.d.



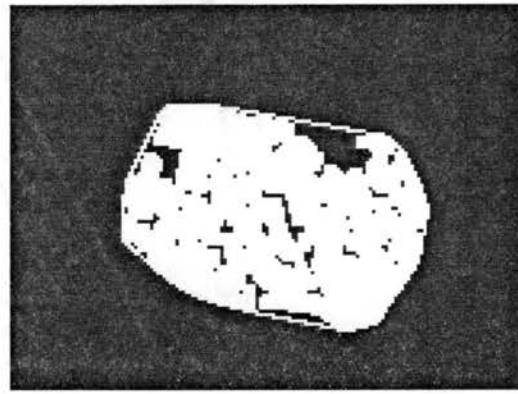
a. Original image.



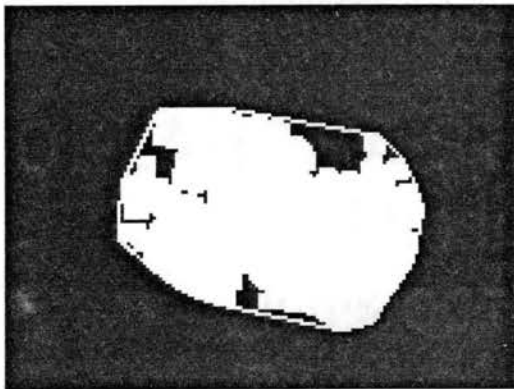
b. After 1st iteration, percent area = 75.07%.



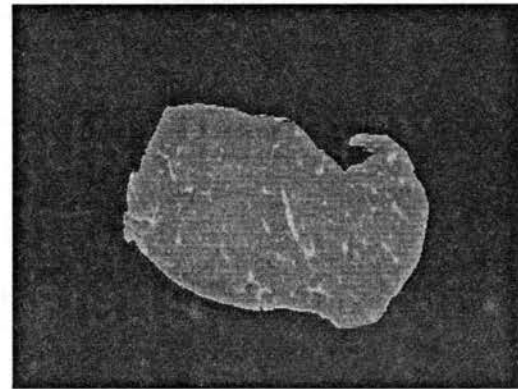
c. After 2nd iteration, percent area = 83.31%.



d. After 3rd iteration, percent area = 86.88%.



e. After 4th iteration; Percent area = 85.03%.



f. Segmented l.d. muscle.

Figure 3.1. Segmentation of longissimus dorsi (l.d.) muscle by adaptive erosion and dilation using convex hull fitting. White line surrounding the muscle is the convex hull.

Oklahoma State University is developing a computer vision system to predict beef quality. The algorithm described by McDonald and Chen (1990) was found to work well for segmenting the l.d. muscle, if threshold value and the parameters of morphological operation (number of iterations and size of the structuring element) were fine-tuned for each image. Threshold value and morphological parameters were fixed in their algorithm, resulting in under-segmentation in some images and over-segmentation in others. In this study, we have developed procedures to automatically optimize threshold value and morphological parameters for each image.

Materials and Methods

Samples

Ribeye rolls from 40 carcasses at 2-d postmortem were obtained from a regional commercial beef processing plant. Criteria for choosing the carcasses included variation in marbling, lean color, and ribeye size. A steak (2.54 cm thick) was removed from the loin end of a ribeye roll from each carcass. Ribeye steaks were vacuum packaged and transported to Oklahoma State University in chilled containers. Steaks were overwrapped in styrofoam trays and chilled at 36°F for 2 to 3 h prior to evaluation. Video images of steaks were acquired at 3-d postmortem. These procedures simulated industry practice in which USDA inspectors evaluate chilled, bloomed ribeyes on carcasses at 2-3 d postmortem. Prior to VIA assessment, two expert graders independently evaluated steaks for USDA marbling score and color score.

Hardware

Images of steaks were acquired using a 550 MHz PC equipped with an Integral Technologies FlashPoint 128 image digitizer. Resolution of the color video camera (A209, MicroImage Video Systems) was 480 x 640 pixels. Output was in RGB format. A dedicated lighting chamber with six 50-watt halogen lamps provided uniform illumination. Steaks were positioned in the camera field-of-view on a matte black surface. The field-of-view was approximately 13.4 x 17.9 cm. Calibrated pixel size was 0.28 by 0.28 mm.

Reference Segmentation

Images from fresh steak samples (n=40) were used in this study (Fig. 3.1a). To evaluate the accuracy of VIA segmentation, an expert grader used graphics software to manually trace the l.d. muscle boundary on a computer monitor. Ambiguous regions surrounding the l.d. muscle challenge even expert graders to identify the ribeye boundary consistently. In order to evaluate the repeatability error or imprecision of manual segmentation, the expert grader traced the l.d. muscle again after two weeks. These traced l.d. muscles were used as the segmentation accuracy reference.

Image Processing

Image processing algorithms were coded in MATLAB 6.0 (Matlab, 2001). Background segmentation was the first operation. The red band from the RGB data gave a well-separated bimodal histogram. Simple thresholding in the red band worked well in distinguishing the background from the steak. The next operation separated lean meat from fat flecks within the ribeye. Thresholding in the green and blue bands discriminated

fat from lean. However, histograms in the green and blue bands displayed a single peak (Fig. 3.2). As a result, finding a critical threshold that was robust for all images was difficult. Thresholding levels were affected by variation among steaks in color of the lean. An adaptive thresholding approach was needed. One of the most common unsupervised segmentation algorithms is clustering (Johnson, 1998). A fuzzy c-means clustering algorithm was used in this study to separate fat from lean.

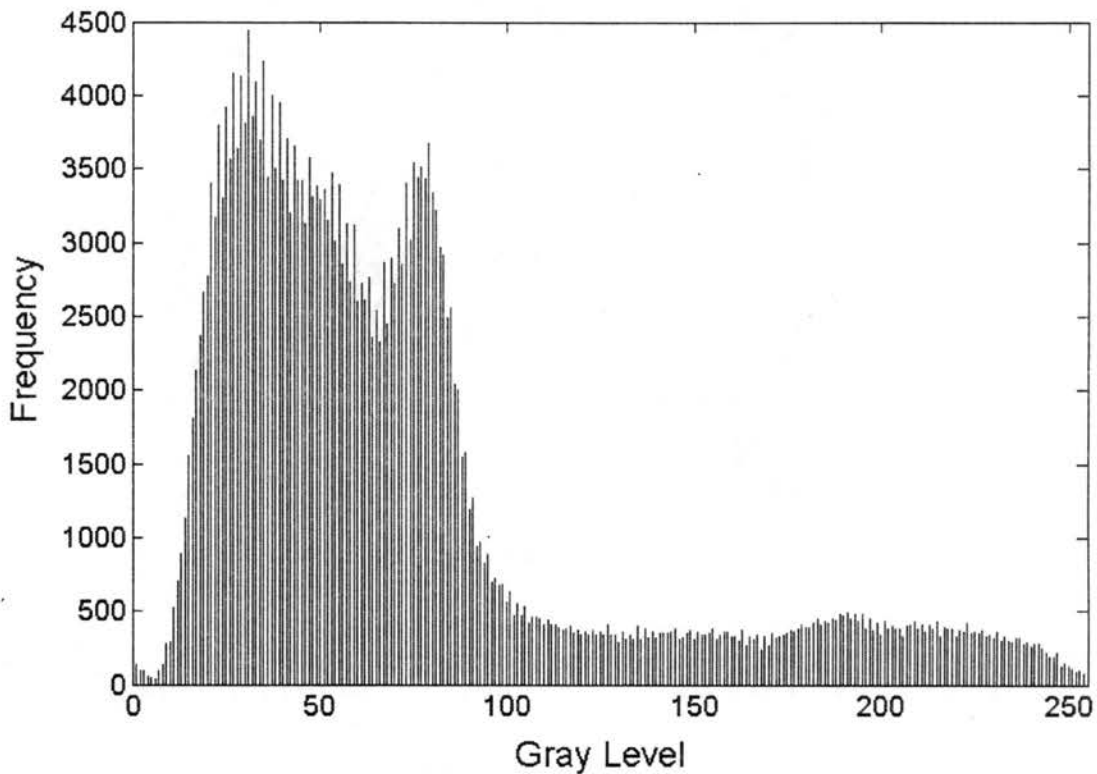


Figure 3.2. Green-band histogram of beef steak image.

Fuzzy C-means Clustering Algorithm

Clustering models can use a range of similarity criteria for grouping data into classes. A suitable model for classifying fat and lean should employ mathematical measures that capture clusters in the way the eye perceives them (Bezdek et al., 1999). The c-means family of clustering models is based on the least squares criterion. C-means

are widely used for batch clustering. Based on optimization, c-means models can be classified as hard, fuzzy, or probabilistic. Simulating human perception, fuzzy c-means (FCM) is suitable for clustering fat and lean meat. The FCM algorithm used in this study is explained below (Bezdek, 1981).

Let us assume that we want to classify the pixels of an image “ I ” into “ C ” classes. Each pixel at position (x,y) has the fuzzy membership value for the i^{th} class, $u_i(x, y)$. The clustering technique is based on minimizing an objective function that accounts for the distance between cluster centers and the data (here in pixel value) within various clusters. The objective function in this case is given by (Bezdek, 1981),

$$J(U, M) = \sum_{x,y} \sum_{i=1}^c (u_i(x, y))^m |d_i(x, y)|^2 \quad (3.1)$$

Where: U = fuzzy c-means class partition,

M = set of cluster centers, and

m = fuzzy exponent, value is greater than 1.

Given a pixel value $I(x, y)$, the measure:

$$d_i(x, y) = |I(x, y) - \mu_i| \quad (3.2)$$

is the distance between the pixel value and the cluster center, μ_i , for the i^{th} class. This distance is weighted in the objective function by the fuzzy membership value for the corresponding pixel. The objective function is iteratively minimized, subject to the following conditions (Bezdek, 1981):

$$\sum_{i=1}^c u_i(x, y) = 1 \quad (3.3)$$

and

$$u_i(x, y) > 0 \quad (3.4)$$

At the start of the iteration scheme, fuzzy membership values are assigned random values and normalized to unity for each pixel. Next, the iteration scheme proceeds, computing the following two quantities in sequence (Bezdek, 1981):

$$u_i(x, y) = 1 / \left[\sum_{x,y} \sum_{l=1}^c \left(\frac{d_l(x, y)}{d_i(x, y)} \right)^{m/(m-1)} \right] \quad (3.5)$$

and

$$\mu_i = \frac{\sum_{x,y} (u_i(x, y))^m I(x, y)}{\sum_{x,y} (u_i(x, y))^m} \quad (3.6)$$

“Convergence” is defined by insignificant change (less than 0.1%) between the two successive values of the objective function. The output of the scheme is the partition, U , of the pixels into C classes. In this study, the imaged steak was grouped into two classes; fat and lean.

Even though FCM is more robust than thresholding, it is computationally intensive. FCM is an iterative procedure. Computation time increases considerably with increase in image size. In order to reduce computation time, a critical factor in real-time applications, the image size was reduced to 120 x 160 pixels by downsampling. Downsampling was performed by selecting every 4th row and column in the original image. Fat and lean were classified by FCM. On all the pixels classified as fat, minimum gray levels on the blue and green bands were determined. These two values were then used to threshold the original image to distinguish the fat areas from lean. Because threshold values are automatically adjusted for every image, this procedure is more robust.

Removal of Intermuscular Fat

Intermuscular fat refers to fat surrounding the l.d. muscle. After classification of fat and lean, intermuscular fat flecks were removed. Blob analysis (Gonzalez and Woods, 1992) was conducted to remove the meat bits that were not touching the l.d. muscle. At this stage, the steak image contained l.d. muscle and some extraneous tissue. In Figure 3.1b, the extraneous tissues are peninsulas touching the periphery of the l.d. muscle. The segmented l.d. muscle is shown in Figure 3.1f.

Morphological Operations

Extraneous tissue can be removed by morphological operations such as erosion and dilation (Serra, 1982). The size and shape of the morphological mask and number of iterations are critical factors for optimal results. Because these parameters vary from sample to sample, fixing the mask size and number of iterations reduces the robustness of the algorithm. These parameters must be adjusted for every sample, based on sample shape. A lower number of iterations of erosion and dilation may not remove all extraneous tissues, whereas an excessive number of iterations will remove portions of l.d. muscle. Optimizing the number of iterations is critical for effective segmentation.

The l.d. muscle is oval (convex) in shape. In our work, the morphological mask was fixed. A 3 x 3 diamond-shaped morphological mask was used. The number of iterations was to be varied to achieve the most compact image. Therefore, a measure of compactness had to be defined. At the end of an erosion operation, blob analysis was conducted to identify the largest component. Separated components were removed, and a dilation operation was performed. An edge map (Gonzalez and Woods, 1992) was

created and a convex hull (described below) was fitted (Fig. 3.1b-1e). The percent area of lean inside the convex hull was used as the compactness measure.

The convex hull is the smallest possible convex polygon that contains all the input points. If each point on the edge map is visualized as a nail protruding from a board, the convex hull is then the shape formed by a rubber band stretched around the nails.

Jarvis's "march" algorithm was implemented to find the convex hull (Cormen et al., 1989). Intuitively, Jarvis's march computes the convex hull by a package-wrapping technique. Input to the algorithm was the edge map of the image resulting from erosion and dilation (Fig. 3.3). For faster implementation, the edge map was first converted from Cartesian to polar coordinates.

Let "n" be the number of edge pixels. The objective of the Jarvis's march algorithm is to determine a sequence, $H = \{p_0, p_1, \dots, p_{h-1}\}$, of vertices of the convex hull that can warp all edge pixels. Figure 3.3 shows the edge map of the image resulting from the morphological operation. All these edge pixels are candidates for the set of vertices of the convex hull. The first vertex chosen, p_0 , is the lowest edge pixel in the edge map (Fig. 3.3). The next vertex, p_1 , has the smallest polar angle with respect to p_0 . Similarly, p_2 has the smallest polar angle with respect to p_1 , and so on. Once we reach the highest vertex, p_k , the next vertex, p_{k+1} , is chosen with the least polar angle with respect to p_k , but from the negative x-axis. The procedure continues, until the first vertex, p_0 , is reached. These vertices can be connected to form the convex hull (Fig. 3.1b).

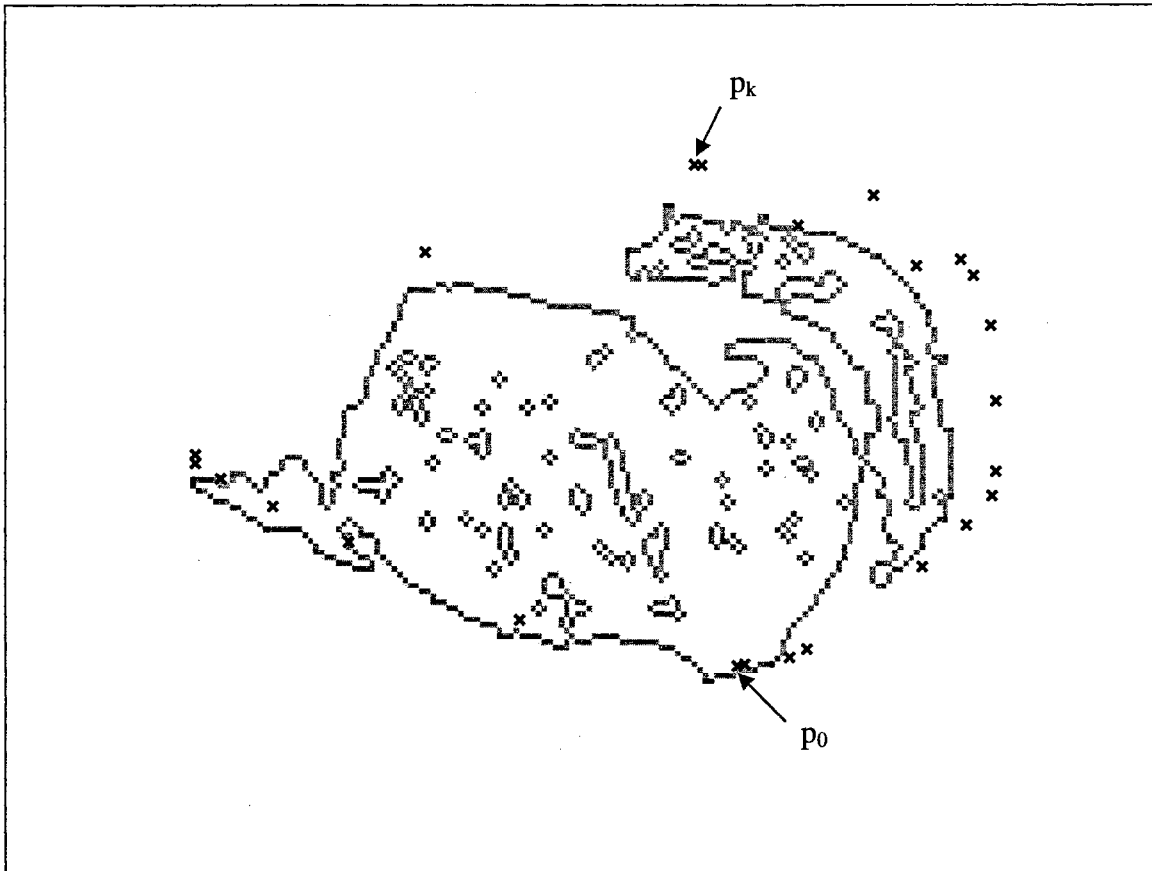


Figure 3.3. Jarvis's march algorithm to determine convex hull. Vertices of convex hull determined by the algorithm are indicated by cross marks.

Percent area of the ribeye inside the convex hull was determined at the end of each iteration (Figs. 3.1b-1e). Iterations were halted when the area of the object was reduced by 70%. This percentage is based on the heuristic that the ribeye occupies at least 70% of the whole steak area. The maximum number of iterations was set at 5, because further iteration split the l.d. muscle into smaller pieces for most images. Fixing a maximum limit on the number of iterations also reduced computation time. After three iterations, holes in the image due to marbling were filled before successive iterations to avoid splitting the l.d. muscle into small components. Percent area of muscle enclosed by the convex hull was analyzed for all iterations. The convex hull in which the maximum

percent area was obtained was selected. The convex hull boundary was filled by a flood-fill operation (Foley et al., 1995). The filled boundary was then combined by logical AND operations with the thresholded original image, in which lean areas were 1's and fat pieces and background were 0's. Holes (due to marbling) were then filled, and an edge map was created. This edge map defined the boundary of the l.d. muscle. The number of iterations of morphological operations was adjusted using the convex hull procedure to obtain the most compact convex l.d. muscle.

Segmentation Accuracy

Classification Error

Percent accuracy of classifying true l.d. muscle pixels as l.d. is given by:

$$LM = \left(\frac{\sum_i^M \sum_j^N AND(R(i, j), S(i, j))}{\sum_i^M \sum_j^N (S(i, j))} \right) 100 \quad (3.7)$$

Where: LM is the l.d. match in %,

R is the reference l.d. image traced by the expert grader,

S is the l.d. image segmented by the algorithm.

The logical operation, AND, is performed pixel-by-pixel. Note that R and S are binary images with l.d. muscle pixels represented as 1 and background pixels represented by 0.

Percent Type I error (l.d. muscle pixel wrongly classified as background pixel) is given by 100-LM. The expert grader traced the l.d. muscle twice to determine the repeatability errors in manual segmentation. As a result, two reference images were generated.

Percent of Type I errors was calculated separately for the two reference images, and the average percent Type I error was recorded.

Percent accuracy of classifying true background pixels as background is given by:

$$BM = \left(\frac{\sum_i^M \sum_j^N NOR(R(i, j), S(i, j))}{\sum_i^M \sum_j^N NOT(S(i, j))} \right) 100 \quad (3.8)$$

Where: BM is the background match in percent. NOR and NOT are logical operations performed pixel-by-pixel. Percent of Type II errors (background pixel wrongly classified as l.d. muscle pixel) is given by 100-BM. Percent classification error was calculated as the average of percent Type I and II errors. Note that classification accuracy may be calculated as (100 - percent classification error).

Average Error Pixel Distance

We propose another measure of segmentation accuracy; Average Error Pixel Distance (AEPD), given by:

$$AEPD = \frac{\sum_i^N \sum_j^M XOR(R(i, j), S(i, j))}{\sum_i^N \sum_j^M P(i, j)} \quad (3.9)$$

where XOR is the “exclusive OR” logical operation and P is the boundary or perimeter of the l.d. muscle traced by the expert grader. The image, P, can be obtained by finding edges of the reference l.d. muscle image, S, using a Sobel operator (Gonzalez and Woods, 1992). XOR produces an output of 1 when there is either a background mismatch or an l.d. muscle mismatch. XOR produces an output of 0 when there is a correct match. Figure 3.4 shows the mismatch between the l.d. muscle segmented by the algorithm (Fig. 3.1f) and the reference image. Sum of mismatch of all pixels in the image is normalized by the length of the boundary of the reference l.d. muscle. The

normalized value, thus, denotes the average distance of the segmented l.d. muscle boundary from the reference l.d. muscle boundary in pixels.

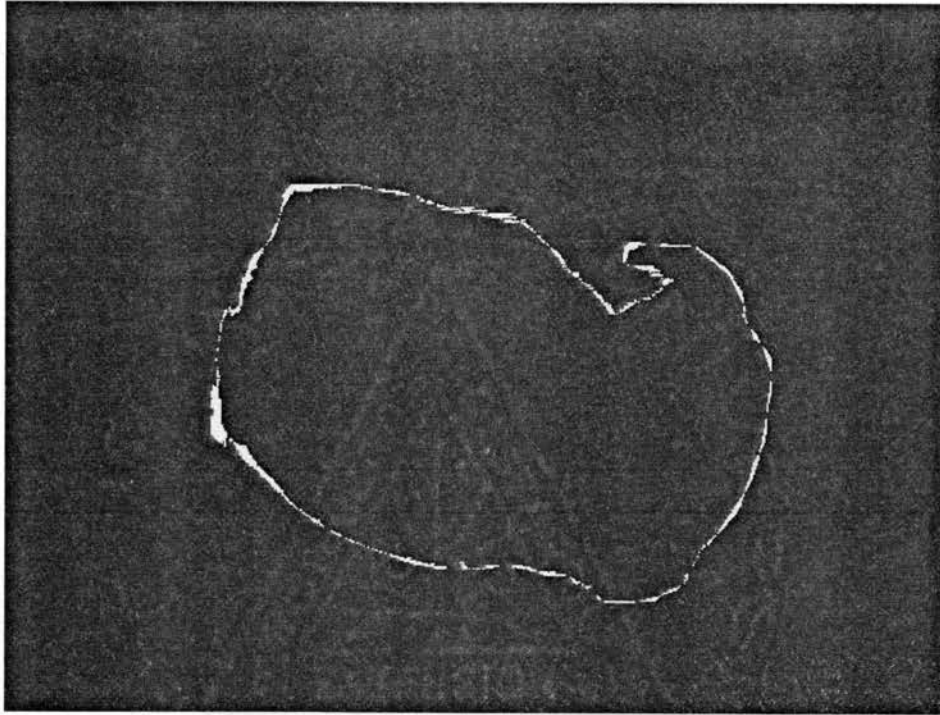


Figure 3.4. Mismatch between the segmented image (Fig. 1f) and the reference image. White regions indicate mismatch, whereas black regions indicate match.

Resolution and Accuracy

In the steak image, the l.d. muscle is large enough to reduce the need for high-resolution (480 x 640) segmentation. By trail-and-error, we found that segmentation accuracy did not drop significantly when the image resolution was reduced by a factor of up to 16. Reducing the resolution increased computational speed. In this study, we have reported results of performing image processing operations on downsampled images with 120 x 160 resolution. For low-resolution images, the morphological mask used was a 2 x 2 square mask, with the maximum number of iterations set at 4. Once the convex hull vertices were determined, the size of the polygon (convex hull) was doubled by scaling

the vertex coordinates by the factor of two. All other image processing operations were the same as those implemented for original-resolution images. Segmentation accuracy and computation times were compared for implementation on both original- and low-resolution images.

Results and Discussion

The adaptive segmentation procedure for a sample image is illustrated in figure 3.1. FCM clustering produced a binarized image with lean = 1, and fat and background = 0. The binary image was eroded and dilated to remove extraneous tissues (Fig. 3.1b). A convex hull was fitted to the image. The white lined polygon surrounding the objects in Figure 3.1b-1e is the convex hull. Percent area of l.d. muscle within the convex hull was calculated as 75.07% after the first iteration. After the second and third iterations, the percent area grew to 83.31 and 86.88%, indicating increased compactness. However, after the fourth iteration, the percent area within the convex hull was reduced to 85%, indicating that some part of l.d. muscle had been eroded. Therefore, the optimum number of iterations for this sample was 3. Based on the criterion of maximum muscle area within the convex hull, the convex hull from the 3rd iteration was selected. The convex hull was filled and combined (logical AND) with the original image that was obtained after FCM clustering. Holes (due to marbling flecks) in the resulting image were filled and multiplied (pointwise, pixel-by-pixel) by the original image (Fig. 3.1a) to define the l.d. muscle (Fig. 3.1f). The number of morphological iterations ranged from 1 to 4 and 1 to 5 for low- and original-resolution images, respectively. In traditional implementation of morphological operations, the number of iterations is fixed. Our algorithm was

designed to select the number of iterations yielding the most compact l.d. muscle area. The result is an algorithm that is more robust.

Percent classification errors, AEPD, and computation time for both low- and original- resolution images for all samples are shown in Table 3.1. To indicate variability, the maximum, minimum, and standard deviation for all measures are given. The average percent classification error and AEPD between the l.d. muscles traced for the same image by the expert grader were 1.19% and 2.3 pixels, respectively, for the 40 images. When a boundary is digitized, an AEPD of at least 1 pixel is inherent. The average repeatability error of the reference method (l.d. muscle traced by the expert grader) with which the algorithm is compared was 1.19% (Table 3.1). One cannot expect any algorithm to perform better than the precision of the reference. Image processing operations on the original-resolution image resulted in an average percent classification error and AEPD between the l.d. muscle segmented by the algorithm and that traced by the expert grader of 1.93% and 3.9 pixels, respectively. Average computation time for the original-resolution images was 16.4 s. The corresponding average percent classification error, AEPD, and computation time for image processing on low-resolution images were 1.97%, 4.4, and 9.8 s (Table 1). Implementing the algorithm on the original-resolution images resulted in minimal improvement in segmentation results (about 0.04% improvement in classification accuracy), but computation time was increased by 40%. Thus, implementing the algorithm on low-resolution images is more suitable for real-time applications. In images No. 2, 8, 25, and 26, classification errors on low-resolution images were lower than those on original-resolution images (Table 3.1). One would expect image processing on original-resolution images would result in lower

Table 3.1. Evaluation of segmentation accuracy and computation time.

ID	Original-Resolution (480x640)			Low-Resolution (120x160)			Expert vs Expert	
	¹ Error (%)	² AEPD	time (s)	Error (%)	AEPD	time (s)	Error (%)	AEPD
1	2.14	3.5	17.66	2.23	3.7	10.58	1.71	3.8
2	2.85	8.2	17.31	1.92	4.6	9.96	1.57	2.6
3	2.09	5.8	16.6	2.11	5.7	9.4	1.43	2.6
4	1.92	5.7	15.19	2.04	5.7	10.18	1.85	3.8
5	1.80	2.7	15.56	2.07	4.2	9.31	1.06	2
6	1.12	1.9	15.88	1.32	2.5	8.95	0.86	1.7
7	1.56	3.4	14.73	1.84	4.1	8.66	1.16	1.9
8	2.19	3.4	16.46	0.88	1.9	9.63	0.82	1.7
9	1.81	2.6	16.69	2.13	2.9	10.15	1.03	1.7
10	1.44	2.6	21.06	1.52	2.7	12.78	1.00	1.9
11	1.13	2.7	15.95	1.50	3.2	8.97	1.35	2.7
12	1.37	3.6	16.23	1.90	5.5	9.55	0.96	1.9
13	1.24	2.8	16.21	1.24	2.8	9.42	1.20	2.2
14	1.68	3.7	16.03	1.78	5.3	9.2	1.13	2.3
15	1.78	3.4	16.75	1.94	3.4	9.68	1.46	2.6
16	1.64	2.9	16.07	1.41	2.7	9.49	1.03	2.1
17	1.31	2.9	16.09	2.26	4	8.63	1.32	2.4
18	1.15	2	16.62	1.32	2.2	9.04	1.09	2
19	1.29	3.1	15.91	1.47	3.1	9.06	1.27	2.4
20	1.93	4	16.23	1.58	3.7	10.07	0.83	1.5
21	3.12	4.8	17.1	2.00	4.2	9.87	2.62	5.4
22	1.35	3.4	15.77	1.46	3.3	9.1	1.03	2
23	1.76	5.5	17.57	1.86	5.5	10.63	1.06	2
24	1.60	4.7	12.76	1.78	4.7	9.39	1.05	1.9
25	2.77	5	16.11	1.46	3.7	9.67	1.05	1.9
26	1.41	2.7	16.68	1.10	2.3	11.1	0.94	1.8
27	1.43	2.4	12.7	1.50	2.5	9.58	0.88	1.8
28	2.13	5	16.45	2.50	7.4	9.11	2.10	3.7
29	2.91	4.8	18.07	1.99	5	11.33	1.02	1.9
30	2.10	5.4	16.98	2.15	5.4	9.5	0.90	1.8
31	1.42	2.6	17.49	1.38	2.7	10.57	0.72	1.4
32	1.22	2.7	17.24	1.35	2.8	9.97	1.22	2.6
33	1.18	2.9	17.01	1.48	4	10.58	1.16	2.5
34	1.41	2.6	15.35	1.05	2.6	8.75	0.63	1.4
35	1.94	3.2	15.03	1.69	3.2	8.73	1.01	1.9
36	1.44	4.9	17.29	1.35	4.9	10.39	1.07	2.4
37	5.86	9.5	16.85	7.12	11.1	9.87	0.88	1.8
38	4.11	6.8	16.66	4.33	7	10.59	2.03	3.9
39	2.59	4.3	16.16	2.23	4	9.73	0.89	1.8
40	2.03	3.6	18	4.49	16.8	11.51	1.06	2
Avg.	1.93	3.9	16.41	1.97	4.4	9.82	1.19	2.3
Max.	5.86	9.5	21.06	7.12	16.8	12.78	2.62	5.4
Min	1.12	1.9	12.7	0.88	1.9	8.63	0.63	1.4
Std.	0.90	1.6	1.36	1.09	2.6	0.87	0.40	0.8

¹Error – Percent classification error. ²AEPD – average error pixel distance, in pixels.

classification errors than on low-resolution images. This contradiction is due to the difference in morphological masks used. A square-shaped mask was better suited to those particular four images than a diamond-shaped mask. The result was higher classification accuracy for those images with low-resolution.

Gao et al. (1995) reported that the classification error of their segmentation algorithms was 15-16%. The adaptive segmentation method presented here had classification errors of less than 2%, representing a significant improvement. Ye (1993) reported that a knowledge-based approach to segmenting the l.d. muscle failed in 20% of the images. Tong et al. (1998) implemented a heuristic approach, but did not report the accuracy of segmentation. It is difficult to develop a knowledge-based or heuristic approach to l.d. segmentation that works on all steaks because of the high biological variability involved. Our approach, using adaptive segmentation, appears to be more robust. Our algorithm was tested with a limited number of samples (n=40). Validation with more samples is recommended.

Conclusion

Segmentation of the *longissimus dorsi* is the critical first step in beef quality grading. In this work, we have implemented and tested an adaptive, robust segmentation method. Threshold values, determined from fuzzy c-means clustering on downsampled images (120 x 160), were used to threshold original-resolution images (480 x 640). Threshold values were adjusted for each image on the basis of FCM clustering, resulting in more accurate classification of fat and lean. Convex hull fitting procedures were used to measure the compactness of the l.d. muscle. The number of iterations of erosion and dilation was selected to achieve the most compact l.d. muscle. The average error pixel

distance and classification errors were 3.9 pixels and 1.93%, respectively, when original-resolution images were used. These algorithms were shown to be sufficiently robust to handle the high biological variability encountered in beef grading. Implementing morphological operations on low-resolution images (120 x 160) and extending the final segmentation results to original resolution lowered classification accuracy minimally (0.04%), but reduced computation time by 40%. The average error pixel distance and classification errors indicating imprecision (repeatability error) of reference expert segmentation were 2.3 pixels and 1.19%, respectively. Classification errors were similar to the imprecision of the reference, indicating strong performance of our adaptive segmentation algorithm. This automated approach showed efficacy and promise for challenging beef-grading tasks.

Acknowledgements

This work was supported by the Research Initiative Program of the Oklahoma State University Food and Agriculture Products Research Center (FAPC). Appreciation is also extended to J.L. Nelson, FAPC Meat Processing Manager, for expert grading assistance.

References

- Belk, K. E., J. D. Tatum, G. Dolezal, B. Morgan, and G. C. Smith. 1996. Meat composition measurement: status of applied research on instrument assessment of composition since completion of the 1994 National Beef Instrument Assessment Planning Symposium. In *Reciprocal Meat Conference Proceedings*, 49: 172-174.
- Belk, K. E., J. A. Scanga, J. D. Tatum, J. W. Wise, and G. C. Smith. 1998. Simulated instrument augmentation of USDA yield grade application to beef carcasses. *Journal of Animal Science* 76(2): 522-527.
- Belk, K. E., J. D. Tatum, G. C. Smith, M. Goldberg, A. M. Wyle, and R. C. Cannell. 2001. Meat imaging system for palatability yield prediction. U.S. Patent No. 5793893.
- Benn, A., D. Barrett-Lennard, and P. J. Hay. 1998. Image analysis for meat. U.S. Patent No. 5793879.
- Bezdek, C. J. 1981. *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York, NY.: Plenum Press.
- Bezdek, C. J., J. Keller, R. Krishnapuram, and N. R. Pal. 1999. *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing*. Boston, MA.: Kluwer Academic Publishers.
- Biju, N. 1998. Beef quality grading with color video image analysis. MS thesis. Stillwater, Oklahoma: Oklahoma State University, Department of Biosystems and Agricultural Engineering.

- Borggaard, C., N. T. Madsen, and H. H. Thodberg. 1996. In-line image analysis in the slaughter industry, illustrated by Beef Carcass Classification. *Meat Science* 43(Supplement 1): 151-163.
- Cormen, T. H., R. L. Leiserson, and R. L. Rivest. 1989. *Introduction to Algorithms*. New York, NY.: McGraw-Hill.
- Foley, J. D., A. V. Dam, S. K. Feiner, and J. F. Hughes. 1995. *Computer Graphics*. 2nd ed. Boston, MA.: Addison Wesley.
- Gao, X., J. Tan, and D. Gerrard. 1995. Image segmentation in 3-dimensional color space. ASAE Paper No. 95-3607. St. Joseph, Mich.: ASAE.
- Goldenberg, A. A., and Z. Lu. 1997. Automation of meat pork grading process. *Computers and Electronics in Agriculture* 16(2): 125-135.
- Gonzalez, R. C., and R. E. Woods. 1992. *Digital Image Processing*. Reading, M.A.: Addison-Wesley Publishing Company.
- Johnson, D. E. 1998. *Applied Multivariate Methods for Data Analysis*. New York, NY.: Duxbury Press.
- Kranzler, G. A., and M. P. Rigney. 1989. Machine vision grading of tree seedlings. In *Proceedings of the 11th International Congress in Agricultural Engineering*. 1883-1888. Rotterdam, The Netherlands.
- Lu, J., and J. Tan. 1998. Application of nonlinear transforms to beef image processing. ASAE Paper No. 98-3016. St. Joseph, Mich.: ASAE.
- Matlab. 2001. *Image Processing Toolbox for Use with Matlab® : User's Guide*. Ver. 3. Natick, MA.: The MathWorks, Inc.

- Mcdonald, T. P., and Y. R. Chen. 1990. Separating connected muscle tissues in images of beef carcass ribeyes. *Transactions of the ASAE* 33(6): 2059-2065.
- NCBA. 2000. The U.S. beef industry: its impact on the American economy. Fact Sheet. Centennial, CO.: National Cattlemen's Beef Association.
- Newman, P. B. D. 1995. Grading and quality control of meat cuts. Great Britain Patent No. 2285126.
- Serra, J. 1982. *Image Analysis And Mathematical Morphology*. New York, NY.: Academic Press.
- Tong, A. K., D. J. Robinson, and T. Liu. 1998. Method and apparatus for using image analysis to determine meat and carcass characteristics. Canadian Patent No. 02263763.
- USDA. 1997. United States standards for grades of carcass beef. Effective date January 31, 1997. Washington, DC.: Livestock and Seed Division of the Agricultural Marketing Service. Available at:
<http://www.ams.usda.gov/lsg/stand/standards/beef-car.pdf>. Accessed on 9 February 2004.
- Ye, J. 1993. A knowledge-based image analysis approach and its application in automatic beef grading. MS thesis. Guelph, Ontario: University of Guelph, Department of Computer Science.

CHAPTER IV
BIOEQUIVALENCE ANALYSIS FOR BEEF QUALITY
GRADING USING COMPUTER VISION

S. Jeyamkondan, G.A. Kranzler, and L. Claypool

Abstract

Assessment of similarity instead of difference between the means of different populations has become an item of growing interest in the area of computer vision for quality inspection of food and agriculture products. This research application presents a classic bioequivalence problem. In this study, the confidence interval approach for solving bioequivalence problems is explained. Then, an application in beef quality grading using computer vision is implemented to demonstrate the use of bioequivalence analysis.

A computer vision system was developed to support automation of beef quality grading. Forty images of ribeye steaks were acquired. After segmenting the *longissimus dorsi* (l.d.) muscle, marbling and color features were extracted to build regression models for predicting marbling and color scores. Quality grade was predicted using another regression model incorporating both marbling and color features. The R^2 values of prediction for color score, marbling score, and quality grade were 0.86, 0.64, and 0.76, respectively. Bioequivalence analysis was conducted. Grades predicted by the computer

vision system were statistically equivalent to the grades assigned by expert graders at $\alpha = 0.05$.

Keywords. Bioequivalence testing, computer vision, beef grading, confidence interval method, equivalence.

Introduction

Traditional statistical analysis is designed to detect the difference in the means of two populations. Researchers typically propose a new technique or method to replace an existing technique. It is the responsibility of the innovator to prove that the new technique is indeed significantly better than the existing technique with $(1-\alpha)100$ confidence, where α is the significance level. Thus, the null hypothesis assumes that the new technique is no better than the existing technique (the two means are equal). The alternate or research hypothesis is that the new technique is better than the existing technique (the two means are different). A $(1-\alpha)100$ confidence interval (CI) is constructed on the difference of the two means. The CI is defined by the lower confidence and upper confidence limits. If the sign, or polarity, of the two limits is different, the null hypothesis is not rejected. If the CI does not cross over zero, the null hypothesis is rejected, and the alternate hypothesis is accepted. So, the statistical analyses are developed to handle the null hypothesis as *equivalence*, and the alternate hypothesis as a *difference*. In the classical statistical sense, failing to reject the null hypothesis does not mean that the equivalence is proved. It is similar to a jury trial in which failing to convict a person does not prove innocence (Steel et al., 1997).

The assessment of similarity instead of difference between the means of different populations has become an issue of growing interest in the area of computer vision for quality inspection of food and agriculture products. Computer vision systems have been developed to automate the grading of various agricultural products such as tree seedlings (Kranzler and Rigney, 1989), apples (Leemans et al. 1998), pork (Goldenberg and Lu, 1997), chicken (Barni et al., 1997), and veal (Denoyelle and Berny, 1999). Before the computer vision system can replace the existing manual grading system, it must prove that the quality grades assigned are statistically equivalent to grades assigned by the manual graders. To prove equivalence, the null and alternate hypotheses must be switched. This example poses a classic problem of bioequivalence.

Bioequivalence problems are commonly encountered in pharmaceutical industries (Munk, 2000). When a drug patent expires, other manufacturers produce new drugs to substitute for the traditional drug. Before substitute drugs are put on the market, manufacturers must prove that the therapeutic effect of the new drug is *equivalent* to that of the existing drug (O'Quigley and Baudoin, 1988). As a result, statistical techniques have been developed to assess the similarity of means (Dannenberg et al., 1994; Holder and Hsuan, 1993; Kinsella, 1989; Lobenberg and Amidon, 2000).

In this study, we have implemented the “confidence interval” method for bioequivalence analysis. The application of beef quality grading using computer vision was implemented to demonstrate the use of the “confidence interval” method for bioequivalence analysis.

Confidence Interval Method

An appropriate null hypothesis for our application is, “Quality traits predicted by the computer vision system are different from those assigned by expert manual graders.”

$$H_0 : \mu_e - \mu_c \neq 0 \quad (4.1)$$

Where: μ_e = mean grade assigned by the expert manual graders,

μ_c = mean grade assigned by the computer vision system.

The alternate or research hypothesis is, “grades are equivalent.”

$$H_a : \mu_e - \mu_c = 0 \quad (4.2)$$

Let $d(i)$ be the difference between the grade assigned by the manual graders, $y_e(i)$, and that assigned by the computer vision system, $y_c(i)$, for the sample, i .

$$d(i) = y_e(i) - y_c(i) \quad (4.3)$$

Let \bar{d} and s_d be the mean and standard deviation, respectively, for the differences for n sample pairs. A confidence interval (CI) is built for the mean of the differences.

Equivalence is claimed if the whole CI is within $\pm c$; where c is the level to be selected by the experimenter. In other words, equivalence is claimed only if:

Lower confidence limit $> -c$, and

Upper confidence limit $< c$.

This procedure is known as the two one-sided test (Kinsella, 1989). Thus, a $(1-2\alpha)100$ two-sided CI is appropriate for this test so that each side is tested separately at the α level of significance (Hsu et al., 1994). The CI is given by

$$CI = \bar{d} \pm t_{\alpha} * s_{\bar{d}} \quad (4.4)$$

Where: $t_{\alpha} = 1 - \alpha$ quantile of a t-distribution with the degrees of freedom, $n-1$,

$$s_{\bar{d}} = \text{standard error} = \frac{s_d}{\sqrt{n}},$$

s_d = standard deviation for the differences for n sample pairs.

Note that in the classical paired t-test, if the CI includes zero, the two means are declared to be *not significantly different* at the 2α level of significance. In order to declare equivalence at the α level of significance, the entire CI must be enclosed within the small range, $\pm c$.

Selection of 'c'

Selection of the value of 'c' can influence the outcome of the study and should be carefully selected and justified. After rigorous studies on bioequivalence of drugs, U.S. Food and Drug Administration (FDA) regulations state that 'c' must be set at 20% of the reference mean because of the high biological variability involved in testing human subjects (Benet 1999).

In this application, the value of CI should be such that a difference of the magnitude of 'c' is of no practical importance in the grading process. Manual grading is subjective. Two expert graders may give different grades to the same sample. Common grading practice requires that the performance of the computer vision system be compared to this subjective manual grading standard with its high variability. The 'c' should be selected, based on the precision of the manual grading system.

Two or more expert graders should assign the grades. Even though the grades assigned by one expert grader may differ from those assigned by another expert grader, these variations will be considered random variation and therefore statistically equivalent. Select a value of c_e (subscript 'e' stands for expert graders) such that the CI at

[(1-2 α)100] is completely contained within $\pm c_e$ to declare equivalence at the α level of significance. Then, 'c' is calculated as:

$$c = kc_e \quad (4.5)$$

Where k is a constant between 1.0 and 2.0, depending on the circumstances. This procedure is based on the concept that the accuracy of the computer vision system in reference to manual graders cannot be greater than the precision of the manual graders.

Beef Quality Grading

In our application, we demonstrate the use of bioequivalence analysis in a computer vision application for beef quality evaluation. Graders assign quality grades from visual appraisal of the *longissimus dorsi* (l.d.) muscle between the 12th and 13th rib. Quality grades are based primarily on marbling levels (abundance and distribution of intramuscular fat on the surface of the l.d. muscle) and on the physiological maturity of the carcass. The USDA defines ten degrees of marbling (Table 4.1) and has published photographs illustrating marbling abundance. With these photographs as reference, graders assign a marbling score. Detailed procedures involved in beef carcass quality grading are published by USDA (USDA, 1997). There is wide interest around the globe in applying computer vision for objective beef grading (Benn et al. 1998; Borggaard, et al., 1996; Newman, 1995; Tong et al., 1998).

A computer vision system was developed to support automation of beef quality grading (Jeyamkondan et al., 2000). Forty images of ribeye steaks were acquired. Two expert graders assigned color scores based on published color standards (Butler et. al., 1980). The standards contain eight photographs to illustrate eight different color classes with

scores from 1 (bleached red) to 8 (very dark red). The expert graders also determined marbling scores and assigned quality grades according to official USDA standards. Marbling scores and USDA quality grades were converted to numerical values following USDA recommendations (Tables 4.1 and 4.2).

Table 4.1. Numerical values for marbling scores.

Marbling Scores*	Numerical Values
Abundant (00-100)	1000-1100
Moderately Abundant (00-99)	900-999
Slightly Abundant (00-99)	800-899
Moderate (00-99)	700-799
Modest (00-99)	600-699
Small (00-99)	500-599
Slight (00-99)	400-499
Traces (00-99)	300-399
Practically Devoid (00-99)	200-299
Devoid (00-99)	100-199

Note: Devoid 100 is same as Practically Devoid 00

* Detailed description of marbling scores is given in USDA (1997). Adapted from Johnson and Dockerty (1990).

The l.d. muscle was segmented from the ribeye using morphological operations. After segmenting marbling by fuzzy c-means clustering (Jeyamkondan et al. 2000), mean and standard deviation of lean color in red, green, and blue bands were calculated. Marbling features extracted include number, area, and perimeter of fat flecks. Stepwise regression models were developed in SAS 8.1 to predict color score, marbling score, and USDA quality grade using extracted features.

Table 4.2. Numerical scale for quality grades.

USDA Quality Grade*	Numerical Scale		
	High (67-99)	Average (34-66)	Low (00-33)
Prime	24	23	22
Choice	21	20	19
Select	18	17	16
Standard	15	14	13
Commercial	11	10	9
Utility	9	8	7

* Detailed description of quality grades is given in USDA (1997).
Adapted from Johnson and Dockerty (1990).

The experimental design was a paired sample design. Expert graders and the computer vision system assigned grades for all samples. Grades were compared for every sample. The objective was to conduct bioequivalence analysis to determine whether the grades assigned by the computer vision system were statistically equivalent to those assigned by the expert graders. In this study, the α value selected was 0.05. In the analysis of equivalence, the significance level will be reported as 0.05. For the analysis of the difference of means, a significance level of 0.10 (2α) will be used.

In this study, we set the value of k at 1.0. Therefore, equivalence is declared if the error of the computer vision system is less than the repeatability error of the manual graders. Note that we compare the vision system error with the variability *between* two expert graders. If the expert grader assigns grade for the same sample a few minutes later, he/she may give a different grade. This difference is called variability *within* the expert grader. We assumed that the *within* variation is much smaller than the *between*

variation, and accordingly, the *between* variation was used as a reference for selecting the value of 'c.'

Marbling Score

The regression model developed for predicting the marbling score using extracted marbling features was:

$$MS = 275.45 + 177.07 * edge - 53.07 * fat - 30.79 * num \quad (4.6)$$

Where: MS = marbling score,

edge = total perimeter of fat flecks per 1000 mm² of l.d. muscle area,

fat = total area of fat flecks per 1000 mm² of l.d. muscle area,

num = number of fat pieces per 1000 mm² of l.d. muscle area.

Note that all features were normalized with respect to l.d. muscle area. The R² value was 0.64. The equivalence analysis for marbling score is given in Table 4.3. Marbling scores assigned by the two expert graders were *not significantly different* at $\alpha = 0.10$. To declare equivalence for marbling scores, c_e must reach 1.35. At this level ($c = c_e = 1.35$), the CI of expert versus computer vision was completely contained within ± 1.35 (Table 4.3) and therefore, the scores assigned by the computer vision system and the experts were *equivalent*. The conceptual meaning is that the marbling score assigned by the computer vision was within ± 1.35 points (which is 13.5% within each marbling level, Table. 4.1) of that given by the expert grader, and we are 95% confident in our procedure. Variability between graders in assigning marbling scores (precision of manual grading for marbling score) was 1.35. Manual grading is the reference standard and computer vision is an instrument to be calibrated. Accuracy of the computer vision

system cannot be greater than the precision of manual grading. Based on this concept, we declare equivalence.

Table 4.3. Equivalence analysis for color and marbling scores and quality grades.

Parameters	Expert 1 vs Expert 2			Expert Average vs CV*		
	Marbling	Color	Grade	Marbling	Color	Grade
Mean of Differences	-0.35	0.84	-0.08	0.00	0.00	0.00
Standard Error of the Differences	0.59	0.12	0.26	0.77	0.07	0.33
Lower Confidence Limit at 90%	-1.35	0.64	-0.34	-1.29	-0.12	-0.33
Upper Confidence Limit at 90%	0.65	1.05	0.18	1.29	0.12	0.33
R ²	0.81	0.66	0.85	0.64	0.86	0.76

*Computer Vision

Color Score

Color features extracted were the mean and standard deviation of the red, green, and blue components of the lean meat. As expected, the color score was highly correlated with the mean value of the red component. The regression equation was:

$$CS = 15.20 - 0.05 * mR \quad (4.7)$$

Where: CS = color score, and

mR = mean value of red component of lean.

The R² value was 0.86. Table 4.3 shows the bioequivalence analysis for color score. The 90% upper and lower confidence limits for expert 1 versus expert 2 did not cross zero.

This result indicates that the color scores assigned by the two expert graders were significantly different at $\alpha = 0.10$. The reason for this discrepancy could be that there are no widely accepted color standards, and therefore graders were not rigorously trained for

assigning color score. Rather, they were well-trained to directly assign carcass maturity. The regression model that used the extracted color features as input generalized well with the average color scores given by the two expert graders. The 90% CI for computer vision versus expert graders was much smaller, and confidence limits crossed zero. Thus, equivalence was declared. The predicted R^2 value (0.86) is much higher than the square of the correlation coefficient between the two expert graders (0.66) (Table 4.3).

Quality Grades

The regression model for predicting the quality grade was:

$$Grade = 12.55 + 0.12 * mR - 0.30 * mB - 0.38 * stdR + 0.41 * stdG + 1.49 * edge \quad (4.8)$$

Where : mB = mean of the blue component of lean meat,

stdR = standard deviation of the red component of lean meat,

stdG = standard deviation of the green component of lean meat.

The R^2 value was 0.76. Table 4.3. shows the bioequivalence analysis for quality grade. For the grades assigned by the two expert graders to be considered statistically equivalent, c_e must reach 0.34. If c is taken as 0.34, the scores assigned by the computer vision system and the experts were *equivalent*. Thus, the grade assigned by the computer vision system was within 0.34 points of that given by the expert graders, and we are 95% confident of our procedure.

Conclusion

The confidence interval method was explained for a bioequivalence problem that is commonly encountered in evaluation of computer visions systems for quality inspection of food and agricultural products. Bioequivalence analysis was conducted for

computer vision applied to beef quality grading to demonstrate use of the confidence interval method. USDA quality grades assigned by the computer vision system were statistically equivalent to those assigned by the expert graders at the significance level of 0.05.

Acknowledgements

Appreciation for funding is extended to the Oklahoma State University Food and Agriculture Products Research Center. The authors express their gratitude to expert graders Dr. G. Dolezol, former Professor of Animal Science, OSU, and B. Schutte, former graduate student in Animal Science, OSU.

References

- Barni, M., V. Cappellini, and A. Mecocci. 1997. Colour-based detection of defects on chicken meat. *Image Vision Computing* 15: 549-556.
- Benet, L. Z. 1999. Understanding bioequivalence testing. *Transplantation Proceedings* 31(Suppl 3A): 7S-9S.
- Benn, A., D. Barrett-Lennard, and P. J. Hay. 1998. Image analysis for meat. U.S. Patent No. 5793879.
- Borggaard, C., N. T. Madsen, and H. H. Thodberg. 1996. In-line image analysis in the slaughter industry, illustrated by Beef Carcass Classification. *Meat Science* 43(Supplement 1): 151-163.
- Butler, J., J. Marchello, E. E. Ray, and R. Shaw. 1980. Standards for Beef Color. Iowa State Extension Publication: AS-515. Ames, IA.: Iowa State University.

- Dannenberg, O., H. Dette, and A. Munk. 1994. An extension of Welch's approximate t-solution to comparative bioequivalence trials. *Biometrika* 81(1): 91-101.
- Denoyelle, C., and F. Berny. 1999. Objective measurement of veal color for classification purposes. *Meat Science* 53: 203-209.
- Goldenberg, A. A., and Z. Lu. 1997. Automation of meat pork grading process. *Computers and Electronics in Agriculture* 16(2): 125-135.
- Holder, D. J., and F. Hsuan. 1993. Moment based criteria for determining bioequivalence. *Biometrika* 80(4): 835-846.
- Hsu, J. C., J. T. G. Hwang, H. Liu, and S. J. Ruberg. 1994. Confidence intervals associated with tests for bioequivalence. *Biometrika* 81(1): 103-114.
- Jeyamkondan, S., N. Ray, G. A. Kranzler, and B. Nisha. 2000. Beef quality grading using machine vision. In *Proceedings of SPIE*, 4203: 91-101. J. A. DeShazer and G. E. Meyer, eds. Boston, MA.: SPIE.
- Johnson, H. K., and T. R. Dockerty. 1990. *Recommended Procedures for Beef Carcass Evaluation and Carcass Contests*. 3rd ed. Savoy, IL.: American Meat Science Association.
- Kinsella, A. 1989. Biostrapping a bioequivalence measure. *Statistician* 38: 175-179.
- Kranzler, G. A., and M. P. Rigney. 1989. Machine vision grading of tree seedlings. In *Proceedings of the 11th International Congress in Agricultural Engineering*. 1883-1888. Rotterdam, The Netherlands.
- Leemans, V., H. Magein, and M. F. Destain. 1998. Defects segmentation on 'golden delicious' apples by using colour machine vision. *Computers and Electronics in Agriculture* 20: 117-130.

- Lobenberg, R., and G. L. Amidon. 2000. Modern bioavailability, bioequivalence and biopharmaceutics classification system. New scientific approaches to international regulatory standards. *European Journal of Pharmaceutics and Biopharmaceutics* 50: 3-12.
- Munk, A. 2000. An unbiased test for the average equivalence problem - the small sample case. *Journal of Statistical Planning Inference* 87: 69-86.
- Newman, P. B. D. 1995. Grading and quality control of meat cuts. Great Britain Patent No. 2285126.
- O'Quigley, J., and C. Baudoin. 1988. General approaches to the problem of bioequivalence. *Statistician* 37: 51-58.
- Steel, R. G. D., J. H. Torrie, and D. A. Dickey. 1997. *Principles and Procedures of Statistics - A Biometric Approach*. 3rd ed. New York, NY.: The McGraw-Hill Companies, Inc.
- Tong, A. K., D. J. Robinson, and T. Liu. 1998. Method and apparatus for using image analysis to determine meat and carcass characteristics. Canadian Patent No. 02263763.
- USDA. 1997. United States standards for grades of carcass beef. Effective date January 31, 1997. Washington, DC.: Livestock and Seed Division of the Agricultural Marketing Service. Available at:
<http://www.ams.usda.gov/lsg/stand/standards/beef-car.pdf>. Accessed on 9 February 2004.

CHAPTER V

**PREDICTING BEEF TENDERNESS FROM
STATISTICAL IMAGE TEXTURAL FEATURES**

S. Jeyamkondan and G.A. Kranzler

Abstract

A computer vision system was developed to predict cooked-beef tenderness. The objective was to predict Warner-Bratzler shear-force tenderness on 14-day aged beef using textural features extracted from 110 fresh beef color images. After ribeye segmentation, images were converted from RGB to CIE LAB color space. Gray-level difference histograms were constructed from each L*, a*, and b* color band, and statistical textural features were extracted. The system predicted shear force with an R² value of 0.50 and correctly classified 79% of samples into two tenderness categories.

Detailed close-up images were also acquired from the final 48 samples. Textural features extracted from these captured more textural information and predicted shear force with a higher R² value of 0.72. Classification into two tenderness groups was 92% correct. Computer vision shows promise for online prediction of cooked-beef tenderness.

Keywords. Computer vision, beef tenderness, Warner-Bratzler shear force, textural features, image processing.

Introduction

The beef industry is the largest food industry in the U.S. (NCBA, 2001). To facilitate marketing, beef grading standards were developed by the USDA to classify carcasses into quality and yield grades. Quality grade categorizes the carcasses on the basis of predicted palatability factors. Beef tenderness is a primary factor determining customer satisfaction. Despite its significance, beef tenderness is not considered in carcass grading. Direct evaluation is absent, because there is currently no suitable method available to the industry for predicting tenderness on-line.

In the 1995 National Beef Quality Audit (Boleman et al., 1998), inadequate tenderness was identified as the second largest quality concern among purveyors, retailers, and restaurants. The National Beef Tenderness Survey (Morgan et al., 1991) showed that consumers experience undesirable toughness in one of every four steaks consumed. Because carcasses are not priced on the basis of tenderness, producers lack incentive to supply a tender product. As a result, consumer preference is not routed back to the producers.

Wulf et al. (1997) investigated the correlation between CIE LAB colorimetric readings (L^* , a^* , and b^*) and beef tenderness and reported that the b^* value showed the highest correlation with Warner-Bratzler shear-force (WBS) tenderness values ($R^2 = 0.14$). Based on this color/tenderness relationship, Colorado State University, in collaboration with Hunter Associates Laboratory, Inc., developed BeefCam[®], a computer vision system to predict beef tenderness.

Vote et al. (2003) evaluated the BeefCam[®] system in four experiments. The R^2 values for predicting WBS tenderness values were 0.17, 0.30, 0.19, and 0.06 for the four

experiments. Wyle et al. (2003) reported that BeefCam[®] predicted WBS values with an R^2 value of 0.18. Of 769 carcasses tested, 13.8% carcasses were “tough.” However, only 51.9% of carcasses were certified as “tender,” which indicates that large numbers of tender carcasses were misclassified as tough. They concluded that further development of the system was necessary for commercial application.

Li et al. (1999a) developed image processing techniques to predict cooked-beef tenderness. In addition to color and marbling features, they used textural features from gray-level co-occurrence matrices of pixel values in RGB color space for predicting sensory tenderness scores. Prediction of sensory tenderness score yielded an R^2 value of 0.70 for the training data set and 0.62 for the testing data set, using 14 factors in a partial least squares model.

Li et al. (2001) used wavelet-based textural features to classify beef carcasses into tender and tough groups. Fifty-nine samples were imaged with a spatial resolution of 480 x 512. Each image was divided into 56 subimages, with 64 x 64 resolution. From 8 extremely tender samples, 45 subimages were selected. Another 45 subimages were selected from 9 extremely tough samples. These 90 subimages were the only images used in the study (Li et al., 1999b). It is not clear why the authors did not use more subimages in this study, when one tough sample can produce 56 subimages. Wavelet textural features ($n=35$) were extracted from each subimage. Fisher’s discriminant model was developed to classify tender and tough samples and was evaluated by cross-validation (leave-one-out technique). They achieved 83.3% correct classification (Li et al., 2001; Li et al., 1999b). Only 17 samples with extreme tenderness scores were used

in this study. If the other samples with intermediate tenderness scores were used, the correct classification rate would obviously be reduced.

In a study supported by National Cattlemen's Beef Association for evaluating instruments for predicting beef tenderness, Wheeler et al. (2002) concluded that the nondestructive methods based on color (BeefCam[®], colorimeter) are not accurate. In summary, there appears to be no satisfactory technique available to predict beef tenderness online. Previous studies by Li et al. (1999a, 1999b, 2001) show that image texture can be used to predict beef tenderness. However, more work is required for implementation.

Recently, we developed a computer vision system to predict USDA quality grade (Biju, 1998). An adaptive thresholding method using a fuzzy c-means algorithm was developed to discriminate fat from lean meat (Jeyamkondan et al., 2000). A robust morphological method was developed to segment the ribeye (Jeyamkondan et al., 2004). Quality grades predicted by the computer vision system were statistically equivalent to those assigned by expert graders (Jeyamkondan et al., 2001).

The objective of this study was to extend our computer vision system to predict shear-force tenderness using image textural features extracted from CIE LAB color space.

Materials & Methods

Samples

Individual beef steaks were analyzed in this study. Fresh steaks (n=110) presenting a broad range of marbling, quality, and tenderness characteristics were

collected from the Oklahoma State University Food and Agriculture Products Center and from regional commercial packing plants (Table 5.1).

Table 5.1. Descriptive statistics of carcass data.

Variable	N	Mean	Std. dev.	Min.	Max.
Shear Force, kg	110	4.30	0.94	2.38	7.60
Hot Carcass Weight, kg	71	323.8	46.8	228.3	474.7
Marbling ^a	110	534.6	150.0	270	1010
Lean Maturity ^b	105	182.5	47.3	130	430
Skeletal Maturity ^b	110	191.2	79.1	120	600
Overall Maturity ^b	105	187.1	58.1	135	515
Quality Grade ^c	110	17.5	2.9	8	24
Ribeye Area, cm ²	101	76.2	12.2	45.8	115.5

^aMarbling: 200=Practically Devoid⁰⁰, 300=Traces⁰⁰, 400=Slight⁰⁰, 500=Small⁰⁰, 600=Modest⁰⁰, etc.

^bMaturity: 100=A⁰⁰, 200=B⁰⁰, etc.

^cUSDA Quality Grade: Ut - = 7, Ut 0 = 8, Ut + = 9, ..., Pr - = 22, Pr 0 = 23, Pr + = 24 (Johnson and Dockerty, 1990).

Computer Vision System

Our computer vision system consisted of a color video camera (A209 MicroImage Video Systems), an image digitizer (FlashPoint 128 Integral Technologies), and a 550 MHz PC and monitor to perform image processing operations.

A dedicated lighting chamber was designed for diffuse and uniform distribution of light. The white interior of an arched cover directed base lighting to a 20x30 cm imaging area (Fig. 5.1). Light was supplied by six 50-watt halogen lamps powered by a feedback controller to stabilize illumination level. The camera was mounted above the lighting chamber, viewing the imaging area through an observation port. A removable

pan with a matte black surface was used to position the steak in the camera field-of-view. Camera resolution was 640 by 480 pixels. Output was in red-green-blue (RGB) format. A calibration grid was employed to quantify pixel size.

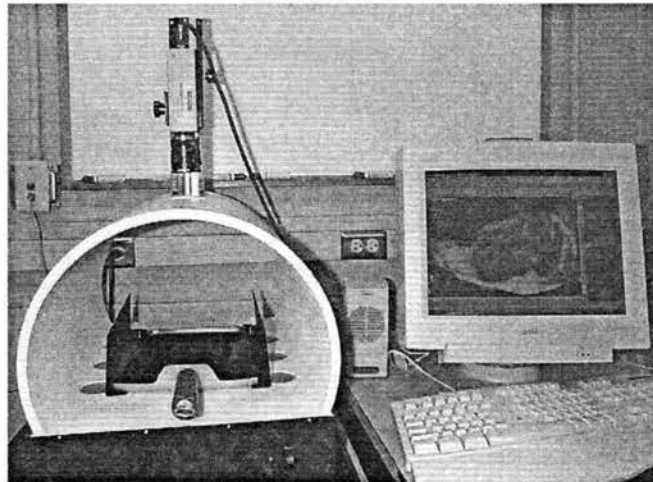


Figure 5.1. Video image analysis system.

An 8-mm lens was mounted on the camera to capture a full-scale image (134 x 179 mm) of the sample steak (Fig. 5.2). For the final 48 samples, close-up images (32 x 43 mm) of the central portion of ribeye were obtained with a 50-mm lens. The close-up images (Fig. 5.3) provided more textural details.

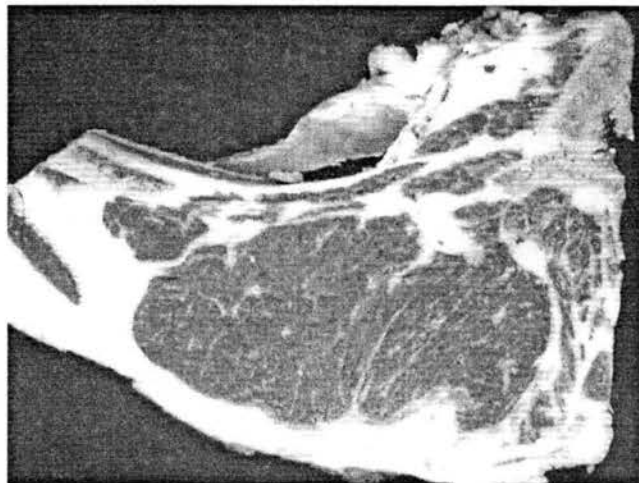


Figure 5.2. Full image of ribeye steak, 8-mm lens.

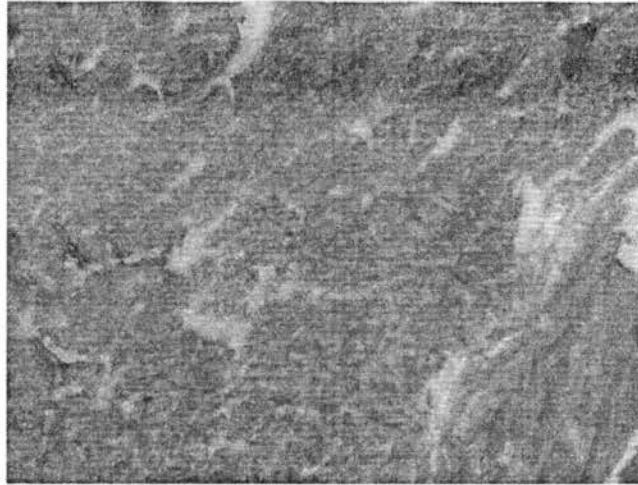


Figure 5.3. Close-up image from central ribeye, 50-mm lens.

Tenderness Reference

After imaging, the steaks were vacuum packaged and aged for 14 days at 1°C. Steaks were then cooked in an impingement oven to an internal temperature of 70°C. Following cooling, six 12-mm diameter core samples were removed parallel to muscle fibers from each steak ribeye. Each core was then sheared using an Instron Universal Testing Machine fitted with a Warner-Bratzler shear-head attachment. The mean of the peak shear-force values from six core samples was taken as the steak tenderness reference. Recommended laboratory guidelines for beef tenderness measurement were followed (AMSA, 1995).

Image Processing

All image processing algorithms were coded in MATLAB 6.0 (The MathWorks, Inc.).

Preprocessing for Full-scale Images

Segmenting the ribeye from the background, intermuscular fat, and extraneous tissue is a critical first step in processing steak images. The ribeye must be isolated for

processing, because shear-force tenderness values are taken from the ribeye, and because experts assign quality grades on the basis of visual appraisal of the ribeye. The segmentation algorithm described by Jeyamkondan et al. (2000, 2004) was used to segment the ribeye.

Color Space Conversion

Color, size, and orientation of patterns influence the perception of texture. The same texture at two different spatial scales will be perceived as two different textures. Color textural features were extracted from ribeyes segmented from full-scale steak images and from the central region of the ribeye taken in close-up images. All images were acquired in RGB format, then transformed to HSI (hue, saturation, intensity) color space. HSI is more compatible with human interpretation. Intensity indicates the brightness of the image, ranging from black to white. Brightness is calculated by averaging values for the R, G, and B color bands. Hue describes a pure color (yellow, orange, or red), whereas saturation describes the degree to which a pure color is diluted by white light. Details for converting RGB to HSI are outlined in Gonzalez and Woods (1992).

RGB images were also transformed to CIE L^* , a^* , b^* (LAB) color space. The CIE LAB is a perceptually uniform color space which mimics the logarithmic response of the eye. Lightness, L^* , ranges from 0 (black) to 100 (white), a^* ranges from -100 (green) to 100 (red), and b^* ranges from -100 (blue) to 100 (yellow).

To convert RGB to CIE LAB, the first step is to transform RGB to CIE XYZ color space. The International Commission on Illumination (CIE, from French title) gives the transformation matrix for converting from RGB to CIE XYZ (Poynton, 1999):

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (5.1)$$

The transformation matrix containing spectral weighting values was derived by the CIE from experiments involving human observers. ‘Y’ is the luminance component of color. Note that the weighting values for luminance sum to one. The above equation assumes the CIE D₆₅ illuminant as a white point. This illuminant has a color temperature of 6504 K and approximates daylight. The CIE defines a normalization procedure to compute *x* and *y* chromaticity coordinates:

$$x = \frac{X}{X+Y+Z}; y = \frac{Y}{X+Y+Z}; z = 1-x-y \quad (5.2)$$

The chromaticity coordinates of D₆₅ are: *x* = 0.3127 and *y* = 0.3291 (Walker, 1996). The luminance, *Y*, is always equal to 100 for a reference white. Procedures for converting CIE XYZ to CIE LAB values are as follows (Connolly and Fleiss, 1997):

$$\begin{aligned} L^* &= 116f\left(\frac{Y}{Y_o}\right) - 16 \\ a^* &= 500\left[f\left(\frac{X}{X_o}\right) - f\left(\frac{Y}{Y_o}\right)\right] \\ b^* &= 200\left[f\left(\frac{Y}{Y_o}\right) - f\left(\frac{Z}{Z_o}\right)\right] \end{aligned} \quad (5.3)$$

Where:

$$f(q) = \sqrt[3]{q} \quad q > 0.008856$$

$$f(q) = 7.787q + \frac{16}{116} \quad q \leq 0.008856$$

*X*_o, *Y*_o, and *Z*_o are the tristimulus values of reference white. For the D₆₅ illuminant, *X*_o, *Y*_o, and *Z*_o values are 95.04, 100, and 108.89, respectively.

Statistical Textural Features

Spatial gray-level co-occurrence matrices (GLCM) comprise the most powerful statistical textural analysis algorithm for extracting second-order textural features (Haralick et al., 1973). They proposed that textural information in an image can be captured in a co-occurrence matrix containing relative frequencies with which a pixel of gray level 'i' occurs with a neighboring pixel of gray level 'j' separated by a relative distance.

Let 'I' be the textural image of size 'K' x 'L'. Each pixel is denoted by $I(k, l)$ with $k=1,2,\dots, K$; and $l=1,2,\dots, L$. Haralick et al. formed the co-occurrence matrix with relative distance defined in polar coordinates (d, θ) , where d is the radial distance and θ is the angle between the two pixels, with respect to the horizontal axis. However, it is easy to describe the co-occurrence matrix using relative distance described in Cartesian coordinates (d_1, d_2) . The probability of observing two pixels with a relative distance (d_1, d_2) in an image with gray levels 'i' and 'j' is given by (Unser, 1986):

$$\begin{aligned} P(y_1 = i, y_2 = j) &= P(i, j) \\ y_1 &= I(k, l) \\ y_2 &= I(k + d_1, l + d_2) \end{aligned} \tag{5.4}$$

The matrix 'P' is called a co-occurrence matrix. Note that the above equation does not depend on absolute indexes (k, l) . If the original image, I, has a dynamic range of N_g , then the size of the co-occurrence matrix, P, is $N_g \times N_g$. Haralick et al. defined fourteen textural features that can be extracted from the co-occurrence matrix. GLCM has been successfully implemented in many applications, including textural analysis of satellite imagery. GLCM is also computationally demanding and consequently too slow for online implementation.

As an alternative, Unser (1986) developed a computationally efficient method called the gray-level difference histogram (GLDH) to calculate second-order textural features. GLDH yields textural features similar to those from GLCM. The key advantage of GLDH is rapid computation. For textural analysis of an image with 256 gray levels, GLDH is not only 64 times faster but also requires 64 times less memory than GLCM. Typically, the variables y_1 and y_2 in Eqn. 5.4 are correlated. Principal component analysis can be applied to produce two new variables (z_1, z_2) that are not correlated and can explain all information in the original variables (y_1, y_2). If the original variables (y_1, y_2) are normalized to have zero mean and unit variance, then the new variables (z_1, z_2) can be calculated by:

$$z_1 = \frac{(y_1 + y_2)}{\sqrt{2}}; z_2 = \frac{(y_1 - y_2)}{\sqrt{2}} \quad (5.5)$$

If the constant ($1/\sqrt{2}$) in the above equation is ignored, the new variables are the sum and difference of the original variables. Thus, the co-occurrence matrix can be replaced by:

$$P(y_1, y_2) = P(z_1, z_2) \quad (5.6)$$

Because z_1 and z_2 are independent (i.e., not correlated), the joint probability function $P(z_1, z_2)$ can be written as a product of the individual probability functions $P(z_1)$ and $P(z_2)$.

$$P(z_1, z_2) = P(z_1)P(z_2) \quad (5.7)$$

Computation is now much simpler.

Based on the above theory, GLDH is implemented as follows (Unser, 1986). First the 'sum' (S) and 'difference' (D) images are calculated as:

$$\begin{aligned}
S(k,l) &= I(k,l) + I(k+d_1, l+d_2) \\
D(k,l) &= I(k,l) - I(k+d_1, l+d_2)
\end{aligned}
\tag{5.8}$$

The dynamic range of the ‘sum’ and ‘difference’ images is twice that of the original image. Histograms of ‘sum’ and ‘difference’ images are calculated as $h_s(i)$ and $h_d(j)$, respectively. The histograms are then normalized to obtain ‘sum’ and ‘difference’ probability functions, as given by:

$$\begin{aligned}
P_s(i) &= \frac{h_s(i)}{KL} & (i = 2, \dots, 2N_g) \\
P_d(j) &= \frac{h_d(j)}{KL} & (j = -N_g + 1, \dots, N_g - 1)
\end{aligned}
\tag{5.9}$$

Seven textural features; energy, correlation, entropy, contrast, homogeneity, cluster shade, and cluster prominence were calculated from the ‘sum’ and ‘difference’ histograms, in addition to mean (μ) and variance of the image. Energy measures textural uniformity, *i.e.*, the repetition of pixel pairs, and is calculated by:

$$Energy = \sum_i (P_s(i))^2 \sum_j (P_d(j))^2$$

Correlation is a measure of gray-level linear dependencies in the image and is calculated by:

$$Correlation = \frac{1}{2} \left\{ \sum_i (i - 2\mu)^2 P_s(i) - \sum_j j^2 P_d(j) \right\}$$

Entropy measures disorder of the image and is calculated by:

$$Entropy = - \sum_i P_s(i) \log(P_s(i)) - \sum_j P_d(j) \log(P_d(j))$$

When the image is not texturally uniform, entropy is large.

Homogeneity, also called an inverse difference moment, is given by:

$$Homogeneity = \sum_j \frac{1}{1+j^2} P_d(j)$$

At constant energy, homogeneity is inversely proportional to contrast. Similarly, at constant contrast, homogeneity is inversely proportional to energy.

Contrast is calculated by:

$$Contrast = \sum_j j^2 P_d(j)$$

Cluster shade and cluster prominence describe the clustering of pixels pairs using third and fourth moments, respectively (Baraldi and Parmiggiani, 1995). They are calculated by (Unser, 1986):

$$Shade = \sum_i (i - 2\mu)^3 P_s(i)$$

$$Prominence = \sum_i (i - 2\mu)^4 P_s(i)$$

The GLDH algorithm can be applied only to gray-scale images and was therefore applied to each color band; red (R), green (G), blue (B), hue (H), saturation (S), intensity (I), lightness (L*), a*, and b*.

Marbling Features

The fuzzy c-means clustering algorithm described in Jeyamkondan et al. (2004) was used to discriminate fat from lean meat. Total numbers of fat flecks, area of fat flecks, and perimeter of fat flecks were calculated.

Statistical Models

Separate stepwise regression models with “maximum R² improvement” as the decision criterion were developed for predicting shear-force tenderness using color

textural and marbling features extracted from full-scale and close-up images. Analysis was performed in SAS 8.1 statistical software.

Results

Prediction of Tenderness from Full-scale Images

The R^2 value for prediction of Warner-Bratzler shear force was 0.50. Samples were classified as 'tough' when shear force equaled or exceeded 10 lbf (44.4 N) and 'tender' when shear force was less than 10 lbf. Seventy nine percent of the steaks were correctly classified into tenderness categories by our computer vision system using full images. The confusion matrix for tenderness classification is given in Table. 5.2.

Table 5.2. Confusion matrix for tenderness classification, full-scale images.

Observed	Percent Classified into	
	Tender	Tough
Tender ^a	84.0	16.0
Tough ^b	31.4	68.6

^aTender: Warner-Bratzler shear force < 10lbf (44.4 N)

^bTough: Warner-Bratzler shear force >= 10lbf (44.4 N)

Prediction of Tenderness from Close-up Images

Textural features from only the CIE LAB color bands L^* and a^* were selected by the stepwise regression procedure. Features selected by the regression model were mean, contrast, cluster shade and cluster prominence of L^* , contrast and entropy of a^* , and number and perimeter of fat flecks. The R^2 value of the regression model for prediction of Warner-Bratzler shear force was 0.72. Overall correct classification rate into two

tenderness categories was 92%. The confusion matrix for classification is shown in Table. 5.3.

Table 5.3. Confusion matrix for tenderness classification, close-up images.

Actual	Percent classified into	
	Tender	Tough
Tender ^a	97.3	2.7
Tough ^b	27.3	72.7

^aTender: Warner-Bratzler shear force < 10lbf (44.4 N)

^bTough: Warner-Bratzler shear force >= 10lbf (44.4 N)

Discussion

Close-up images captured more textural information and therefore explained more variation in beef tenderness than full-scale images. Better prediction by close-up images suggested that the central portion of the ribeye was representative of the entire ribeye. Because ribeye segmentation was not required, close-up images simplified the program and reduced computation time.

Li et al. (1999a) extracted textural features using GLCM and gray-level pixel value run-length features in RGB color space. They developed a PLS model with 14 factors that predicted sensory tenderness with R^2 values of 0.70 and 0.62 for training set and testing set, respectively. We have implemented GDLH, a less complex algorithm in CIE LAB color space that is fast enough for on-line application. The eight-feature regression model predicted Warner-Bratzler shear force with an R^2 value of 0.72. The stepwise regression models selected features extracted from the L^* and a^* bands, in addition to two marbling features. Wulf et al. (1997) reported that b^* values showed the

highest correlation with shear force ($R^2 = 0.14$). In our study, none of the textural features extracted from b* band was targeted by the stepwise selection procedure. Wulf et al. (1997) used a colorimeter to obtain average L*a*b* values from each steak. In our study, RGB images was converted to CIE LAB images, and therefore L*a*b* values were available at each pixel location in the image. More accurate results in this study may be attributable to the fact that tenderness may be related to textural distribution of L*a*b* values, rather than average L*a*b* values.

Enzymatic action during aging tenderizes beef. The beef industry, however, is interested in predicting tenderness of the carcass at the time of slaughter so that carcasses can be marketed on the basis of their true value related to customer preference. In this study, we demonstrated a computer vision system that can predict aged-beef tenderness from fresh-beef images. Images were captured before aging and models were developed to predict tenderness after 14 days of aging.

Conclusion

Textural features in CIE LAB color space were extracted using a gray-level difference histogram algorithm. Close-up images contained more textural information and predicted tenderness more accurately than full-scale images. Color textural features extracted from close-up images predicted Warner-Bratzler shear-force tenderness values with an R^2 value of 0.72. Overall correct classification into tenderness categories of 'tender' and 'tough' was 92%. Implementing the computer vision system with close-up images not only increased prediction accuracy, but also reduced computation time. Computer vision shows promise as a nondestructive method of predicting beef tenderness that can be implemented online.

Acknowledgements

The authors thank the Food and Agriculture Products Research Initiative Program (FRIP) 2000, Oklahoma State University for funding this study. Appreciation is expressed to Drs. B. Morgan and C. Brooks, OSU Animal Science, for providing technical expertise on beef grading and tenderness. The authors also thank Laura Locke, Animal Science Graduate Research Assistant for generating shear-force data.

References

- AMSA. 1995. *Research guidelines for cookery, sensory evaluation, and instrumental tenderness measurements of fresh meat*. Savoy, IL.: National Live Stock & Meat Board, American Meat Science Association.
- Baraldi, A., and F. Parmiggiani. 1995. An investigation of the textural characteristics associated with gray level cooccurrence matrix statistical parameters. *IEEE Transactions on Geoscience and Remote Sensing* 33(2): 293-304.
- Biju, N. 1998. Beef quality grading with color video image analysis. MS thesis. Stillwater, Oklahoma: Oklahoma State University, Department of Biosystems and Agricultural Engineering.
- Boleman, S. L., S. J. Boleman, W. W. Morgan, D. S. Hale, D. B. Griffin, J. W. Savell, R. P. Ames, M. T. Smith, J. D. Tatum, T. G. Field, G. C. Smith, B. A. Gardner, J. B. Morgan, S. L. Northcutt, H. G. Dolezal, D. R. Gill, and F. K. Ray. 1998. National Beef Quality Audit-1995: survey of producer-related defects and carcass quality and quantity attributes. *Journal of Animal Science* 76(1): 96-103.

- Connolly, C., and T. Fleiss. 1997. A study of efficiency and accuracy in the transformation from RGB to CIELAB color space. *Image Processing, IEEE Transactions on* 6(7): 1046-1048.
- Gonzalez, R. C., and R. E. Woods. 1992. *Digital Image Processing*. Reading, M.A.: Addison-Wesley Publishing Company.
- Haralick, R. M., K. Shanmugam, and I. Dinstein. 1973. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics* SMC-3(6): 610-621.
- Jeyamkondan, S., N. Ray, G. A. Kranzler, and B. Nisha. 2000. Beef quality grading using machine vision. In *Proceedings of SPIE*, 4203: 91-101. J. A. DeShazer and G. E. Meyer, eds. Boston, MA.: SPIE.
- Jeyamkondan, S., and G. A. Kranzler. 2001. Bioequivalence analysis for beef quality grading with computer vision. In *2001 Oklahoma Section Meeting of the American Society of Agricultural Engineers*. Stillwater, OK.
- Jeyamkondan, S., N. Ray, G. A. Kranzler, and S. Acton. 2004. Computer vision segmentation of the longissimus dorsi for beef quality grading. *Accepted for publication in the Transactions of the ASAE*.
- Johnson, H. K., and T. R. Dockerty. 1990. *Recommended Procedures for Beef Carcass Evaluation and Carcass Contests*. 3rd ed. Savoy, IL.: American Meat Science Association.
- Li, J., J. Tan, F. A. Martz, and H. Heymann. 1999a. Image texture features as indicators of beef tenderness. *Meat Science* 53(1): 17-22.

- Li, J., J. Tan, and P. Shatadal. 1999b. Discrimination of beef images by textural features. ASAE Paper No. 99-3158. St. Joseph, Mich.: ASAE.
- Li, J., J. Tan, and P. Shatadal. 2001. Classification of tough and tender beef by image texture analysis. *Meat Science* 57(4): 341-346.
- Morgan, J. B., J. W. Savell, D. S. Hale, R. K. Miller, D. B. Griffin, H. R. Cross, and S. D. Shackelford. 1991. National beef tenderness survey. *Journal of Animal Science* 69(8): 3274-3283.
- NCBA. 2001. The U.S. beef industry: its impact on the American economy. Fact Sheet. Centennial, CO.: National Cattlemen's Beef Association. Available at: http://www.beef.org/dsp/dsp_content.cfm?locationId=710&contentType=1&contentId=252. Accessed on 11 February 2004.
- Poynton, C. 1999. Frequently asked questions about color. Available at: <http://www.poynton.com/PDFs/ColorFAQ.pdf>. Accessed on 15 February 2004.
- Unser, M. 1986. Sum and difference histograms for texture classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 8(1): 118-125.
- Vote, D. J., K. E. Belk, J. D. Tatum, J. A. Scanga, and G. C. Smith. 2003. Online prediction of beef tenderness using a computer vision system equipped with a BeefCam module. *Journal of Animal Science* 81(2): 457-465.
- Walker, J. 1996. Colour rendering of spectra. Switcherland: Fourmilab. Available at: <http://www.fourmilab.ch/documents/specrend/>. Accessed on 15 February 2004.

- Wheeler, T. L., D. Vote, J. M. Leheska, S. D. Shackelford, K. E. Belk, D. M. Wulf, B. L. Gwartney, and M. Koochmarai. 2002. The efficacy of three objective systems for identifying beef cuts that can be guaranteed tender. *Journal of Animal Science* 80(12): 3315-3327.
- Wulf, D. M., S. F. O'connor, J. D. Tatum, and G. C. Smith. 1997. Using objective measures of muscle color to predict beef longissimus tenderness. *Journal of Animal Science* 75(3): 684-692.
- Wyle, A. M., D. J. Vote, D. L. Roeber, R. C. Cannell, K. E. Belk, J. A. Scanga, M. Goldberg, J. D. Tatum, and G. C. Smith. 2003. Effectiveness of the SmartMV prototype BeefCam System to sort beef carcasses into expected palatability groups. *Journal of Animal Science* 81(2): 441-448.

CHAPTER VI
TEXTURE ANALYSIS OF BEEF IMAGES
USING THE GABOR FILTER

S. Jeyamkondan, G.A. Kranzler, and G. Fan

Abstract

Tenderness is a critical factor in consumer perception of beef palatability. We developed algorithms to extract textural features from beef digital images to predict aged, cooked-beef tenderness that can be easily incorporated into existing computer vision systems.

Images of fresh beef steaks ($N = 186$) were acquired using a computer vision system. We used a Gabor filter to extract textural features. The Gabor filter is a bank of band-pass filters designed in the spatial domain. The filter kernel is a Gaussian modulated by a complex exponential. A powerful property of the Gabor filter that makes it suitable for this application is that the filter is localized in both the spatial and frequency domains. A four-scale and six-orientation Gabor filter consists of a bank of 24 band-pass filters. A digital image processed through this filter is decomposed into 24 images, each with a distinct frequency band and orientation. The mean and standard deviation of these decomposed images were calculated.

A canonical discriminant model was developed to reduce 48 textural features to 2 canonical features. The model classified the steaks into three tenderness categories with 79.6% accuracy using the leave-one-out cross validation method.

Keywords. Beef tenderness, texture features, computer vision, image analysis, Gabor filter, canonical discriminant analysis.

Introduction

The U.S. beef industry forms the largest part of the nation's food and fiber industry, with a retail value of \$57 billion in 2001 (NCBA, 2003). In Oklahoma, cattle sales represent 53.1% of all agricultural marketing (OBIC, 2003). USDA developed beef grading standards to facilitate marketing (USDA, 1997). USDA quality grade depends on marbling as a major factor influencing palatability. However, the literature shows contradictory results regarding the influence of marbling on tenderness (Jeremiah, 1996). The National Beef Tenderness Survey-1998 revealed that USDA quality grade has little or no effect on consumer sensory evaluation or shear-force tenderness values (Brooks et al., 2000).

Tenderness is a critical factor in consumer perception of beef palatability. National Beef Quality Audits conducted in 1991 (Lorenzen et al. 1993), 1995 (Boleman et al., 1998), and 2000 (McKenna et al., 2002) identified inadequate tenderness as one of the largest quality concerns in the beef industry. Despite its significance as a quality indicator, tenderness is not a factor affecting product value for producers and packers, because the USDA quality grading system does not currently incorporate tenderness. Direct evaluation is absent, because there is currently no industry accepted method available for predicting tenderness online.

Currently, computer vision or video image analysis (VIA) systems are available to predict beef yield and quality grade (Benn et al., 1988; Tong et al., 1998; Belk et al., 2001). VIA systems capture images of fresh ribeye and use them to predict yield and quality grades. The USDA has approved the use of VIA systems for grader-assisted determination of beef yield grades (USDA, 2001). However, VIA use for quality grade assignment awaits approval.

In this study, we have developed algorithms to extract textural features from beef images to predict aged, cooked-beef tenderness that can be easily incorporated into existing computer vision systems. Implementation of the Gabor filter algorithm is explained.

Materials and Methods

Samples

Beef strip loin steak samples (n=186) were collected from regional packing plants. Quality grades of the tested carcasses were generated to reflect the proportion of quality grades typically found in U.S. packing plants (i.e., 2.3% Prime, 41.2% Choice, 43.3% Select, and 3.4% Standard) as reported by the recent National Beef Quality Audit – 2000 (McKenna et al., 2002). Beef samples were prepared for imaging at the Oklahoma State University Food and Agricultural Products Research Center. One-inch steaks were cut, individually identified, and allowed to “bloom” for 30 minutes prior to imaging.

Computer Vision System

The computer vision system consisted of a color video camera (A209 MicroImage Video Systems), an image digitizer (Flash Point 128 Integral Technologies) and a 1.0 GHz PC. Images of the beef steaks were captured in a diffuse, uniform lighting chamber. Light was provided by six 50-watt halogen lamps powered by a stabilizing feedback controller. The camera was mounted above the lighting chamber, viewing a 32x43 mm imaging area through an observation port. Camera resolution was 640x480 pixels. Pixel size was 67 x 67 μm . Output was in the red-green-blue (RGB) format.

Tenderness Reference Measurement

After imaging, samples were vacuum-packaged and aged for 14 days. Steaks were then cooked in an impingement oven to an internal temperature of 70° C. After cooling, six 12-mm diameter cores were sampled randomly from each steak ribeye. Cores were removed parallel to the muscle fibers. Force required to shear the core was recorded using a Warner-Bratzler attachment to an Instron Universal Testing Machine. Peak shear-force values from the six replicate cores were averaged to obtain the shear-force tenderness reference value (AMSA, 1995). Higher shear-force values indicated “tougher” beef. Steaks were identified as “tender,” “medium,” and “tough” if the shear-force values were less than 8.5 lbf (37.7 N), between 8.5 and 10 lbf, and greater than 10 lbf (44.4 N), respectively.

Image Texture

Texture is related to the frequency content of an image. An image from the spatial domain can be converted to the frequency domain using the Fourier transform

(FT). Energy concentrated at the center of the frequency spectrum indicates low-frequency content (textural coarseness), whereas energy from regions far from the center of the frequency spectrum indicates high-frequency content (textural fineness). The spatial domain (original image) contains complete spatial information, but no frequency information. The FT identifies all spectral components of an image, however it does not contain spatial information. The FT analyzes the image globally, not locally. In order to obtain spatial localization of frequency components, the image must be analyzed locally. To meet this requirement, the short-time Fourier transform (STFT) was introduced. An image was spatially localized by multiplying the image by a spatial window. The windowed image is then Fourier-transformed to obtain frequency components of that spatially localized image. Various windows including rectangular, triangular, and Gaussian can be used. When the window is a Gaussian, the STFT is also known as the Gabor Transform, from its inventor, Dennis Gabor (Misiti et al., 1996). The Gabor transform is implemented with a bank of band-pass digital filters. As a result, the name “Gabor filter” is commonly used in the literature.

Digital Filter

A digital filter can be used to separate certain frequency content in an image. For instance, a high-pass filter attenuates low-frequency content and passes high-frequency content. The converse is true for a low-pass filter. A band-pass filter passes only a specific band of frequencies. A digital filter is called a linear filter when it satisfies the following conditions (Oppenheim and Schaffer, 1991):

$$H(cf+dg) = cH(f)+dH(g) \quad (6.1)$$

Where: $H(f)$ is the output of a linear filter on an image, f ,

f and g are images,

c and d are constants.

A space-invariant filter means that the filter characteristics do not vary with the location of the image. A linear space-invariant (LSI) filter can be completely characterized by its impulse response (Oppenheim and Schaffer, 1999). An impulse can be represented digitally by a matrix with a value 1 at the center and 0 at all other locations. The output image of the LSI filter is given by the convolution of the impulse response of the filter with the input image:

$$g = h ** f \quad (6.2)$$

Where: g is the output image from the filter,

f is the input image to the filter,

h is the impulse response of the filter,

** is the two-dimensional convolution operation.

The concept of the convolution may be seen in an image or signal processing textbook such as Oppenheim and Schaffer (1991). Conducted in the spatial domain, the convolution operation is computationally intensive. Convolution in a spatial domain is equivalent to multiplication in the frequency domain. As a result, convolution is usually implemented in the frequency domain. An FT of the impulse response reveals the frequency response of the filter. The frequency response indicates how the filter attenuates or amplifies frequency content in the input image. Equation 6.2 in spatial domain is equivalent to the following equation in the frequency domain:

$$G(u,v) = H(u,v)F(u,v) \quad (6.3)$$

Where: G = FT of the output image, g ,

H = frequency response of the filter, which is FT of impulse response, h ,

F = FT of the input image, f .

If we take the inverse Fourier transform (IFT) of the output (G), we obtain the filtered image, g .

Gabor Filter

The algorithm for the Gabor filter described in Manjunath and Ma (1996) was implemented in Matlab 6.1. (Mathworks, Natick, MA). Band-pass filters are designed in the spatial domain. The kernel function of an FT is a complex exponential, whereas the kernel function of a Gabor filter is a Gaussian function modulated by a complex exponential. A Gaussian is a smooth bell-shaped curve. The width of a Gaussian can be defined by the parameter, 'variance.' Because the FT of the Gaussian is also a Gaussian (Oppenheim and Schaffer, 1999), the window is smooth in both the spatial and frequency domains. When the Gaussian is rotated in the spatial domain, the Gaussian in the frequency domain also rotates. Variances of Gaussians in the spatial and frequency domains are inversely related (Oppenheim and Schaffer, 1999).

The frequency responses of a bank of band-pass filters are designed with the following objectives:

- Combined frequency response of all band-pass filters must cover either one-half of the frequency spectrum. The other half has redundant information, because the FT of a real image is conjugate-symmetric.

- There should be minimal overlap between the frequency response of a given band-pass filter and that of adjacent band-pass filters.

Figure 6.1 shows frequency responses of 24 band-pass filters. Contours indicate the half-peak magnitude of the frequency response. Note that the shape of the frequency response is a Gaussian and that the half-peak magnitude contour of a band-pass filter just touches the contours of other filters.

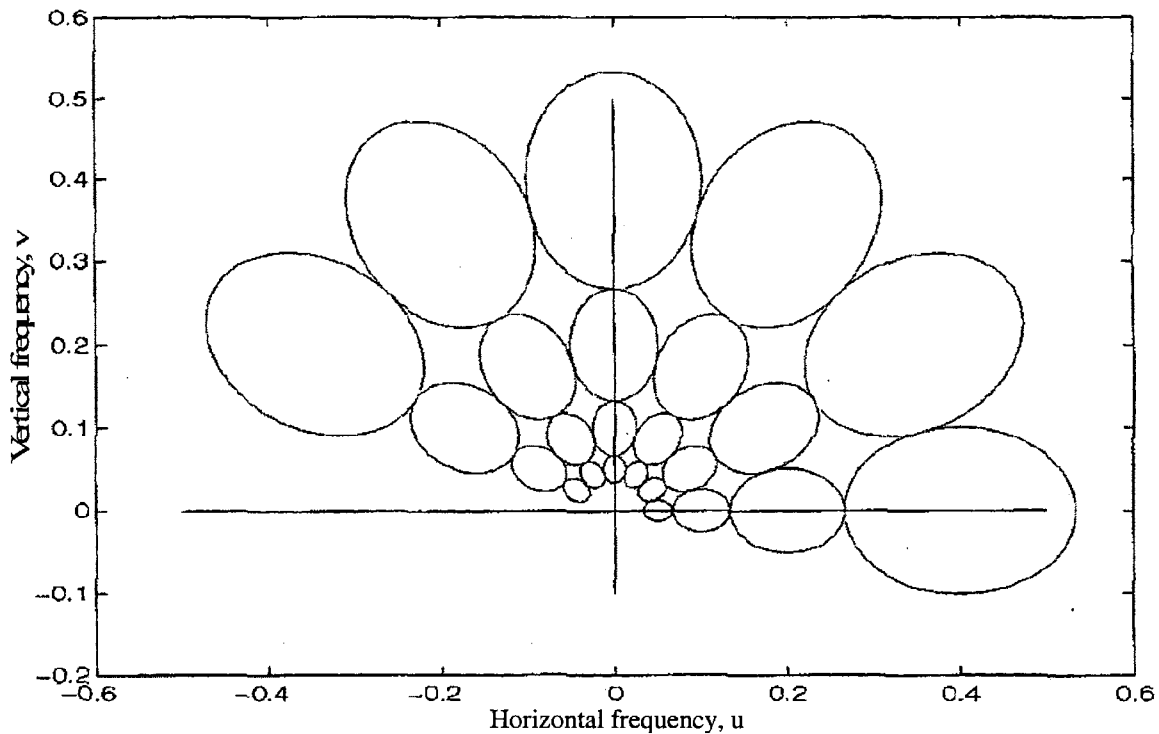


Figure 6.1. Gabor filter design (Adapted from Manjunath and Ma, 1996).

Gabor filter design requires four parameters: number of scales (S), number of orientations (K), center frequency of the Gaussian of the coarsest scale (the Gaussian with the smallest contour size) along the u -axis (U_l), and center frequency of the Gaussian of the coarsest scale (the Gaussian with the smallest contour size) along u -axis (U_h). Parameters for the Gabor filter shown in Figure 6.1 are: $S=4$, $K=6$, $U_l=0.05$, and $U_h=0.4$.

Variances of Gaussians in the frequency domain are determined by:

$$\sigma_u = \frac{(a-1)U_h}{(a+1)\sqrt{2\ln 2}},$$

$$\sigma_v = \frac{\tan\left(\frac{\pi}{2K}\right)\left(U_h - 2\ln\left(\frac{2\sigma_u^2}{U_h}\right)\right)}{\sqrt{2\ln 2 - \frac{(2\ln 2)^2 \sigma_u^2}{U_h^2}}}$$

Where: σ_u = variance of the Gaussian kernel along u-axis,

σ_v = variance of the Gaussian kernel along v-axis,

$$a = \left(\frac{U_h}{U_l}\right)^{\frac{1}{s-1}}.$$

Then, the variances of the Gaussian in the spatial domain are calculated by:

$$\sigma_x = \frac{1}{2\pi\sigma_u}, \quad \sigma_y = \frac{1}{2\pi\sigma_v}$$

Where: σ_x = variance of the Gaussian kernel along x-axis,

σ_y =variance of the Gaussian kernel along y-axis,

The Gabor filter is then designed in the spatial domain by:

$$g(x, y) = a^{-m} \left(\frac{1}{2\pi\sigma_x\sigma_y} \right) \exp\left(-\frac{1}{2} \left(\frac{(x')^2}{\sigma_x^2} + \frac{(y')^2}{\sigma_y^2} \right) + j2\pi x'U_h \right) \quad (6.4)$$

Where: x' and y' are oriented/rotated coordinates

$$x' = a^{-m} \left(x \cos\left(\frac{n\pi}{K}\right) + y \sin\left(\frac{n\pi}{K}\right) \right),$$

$$y' = a^{-m} \left(-x \sin\left(\frac{n\pi}{K}\right) + y \cos\left(\frac{n\pi}{K}\right) \right),$$

$$m = 0, 1, \dots, S-1,$$

$$n = 0, 1, \dots, K-1.$$

To illustrate Gabor filter design, a Gabor filter of size 128 x 128 with three scales and three orientations is selected. Figure 6.1 shows the Gabor filters for an orientation ($n = 0$; horizontal in spatial; vertical in frequency domain) and three scales. The first row in Figure 6.1 represents the coarsest scale, the last row represents the finest scale, and the middle row represents the medium scale. The first and second columns represent real and imaginary parts, respectively, of the impulse response of the band-pass filter. The third column represents the frequency response of the band-pass filter.

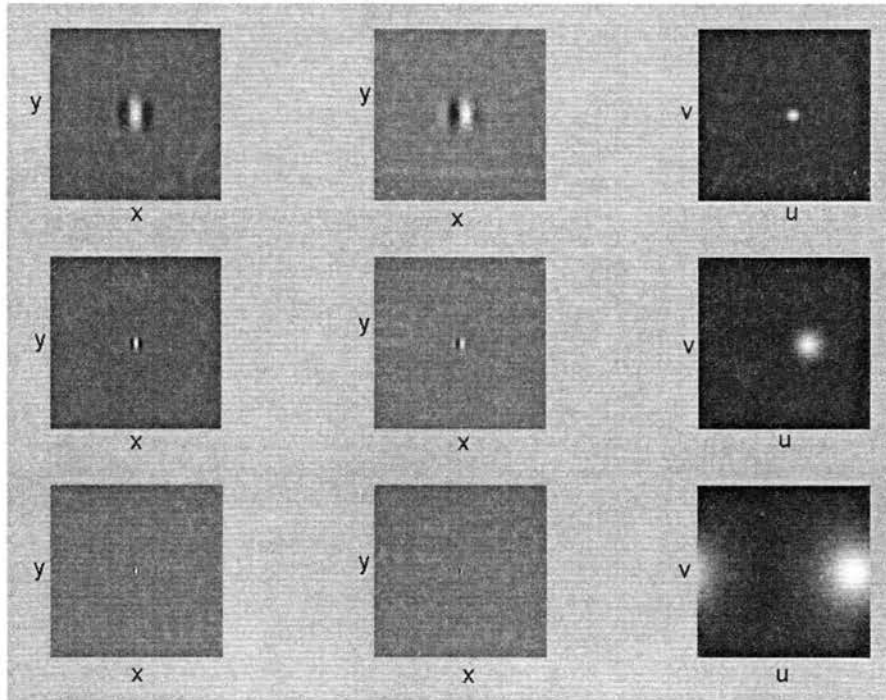


Figure 6.2. Impulse and frequency response of Gabor filter ($K=3$; $S=3$; $n=0$; $m=0,1,2$).

The following observations can be made from Figure 6.2.

1. Both impulse and frequency responses are Gaussian shaped.

2. Impulse response is a Gaussian modulated by an exponential. The real part is the Gaussian modulated by a cosine, whereas the imaginary part is the Gaussian modulated by a sine.
3. Frequency response is also a Gaussian, but shifted by the amount equal to the frequency of the exponential which modulated the impulse response.
4. Variance of Gaussian in the frequency domain is inversely related to the variance of the Gaussian in the spatial domain.
5. Variance of the Gaussian in the spatial domain for coarse scale (row 1) is largest, whereas that for fine scale (row 3) is smallest.
6. Variance of the Gaussian in the frequency domain for the coarse scale (row 1) is smallest, and the center of the Gaussian is closer to the center of the FT image. Therefore, this filter captures the low-frequency content of the image.
7. Variance of the Gaussian in the frequency domain increases by log base 2, as the scale or the central frequency of the Gaussian increases (from row 1 to row 3). This response simulates human vision, which is more sensitive to low-frequency content than high-frequency content. Therefore, the bandwidth of the band-pass filter increases as the central frequency increases.

Column 3 in Figure 6.2 is the image display of the frequency response of the Gabor filters. The gray-level corresponds to the amplitude of the frequency response.

Another method of visualization is a three-dimensional mesh plot. Figures 6.3-6.5 show the frequency response for coarse, medium, and fine scale. These frequency responses are band-pass in nature. The response shape is Gaussian, and the surface is smooth.

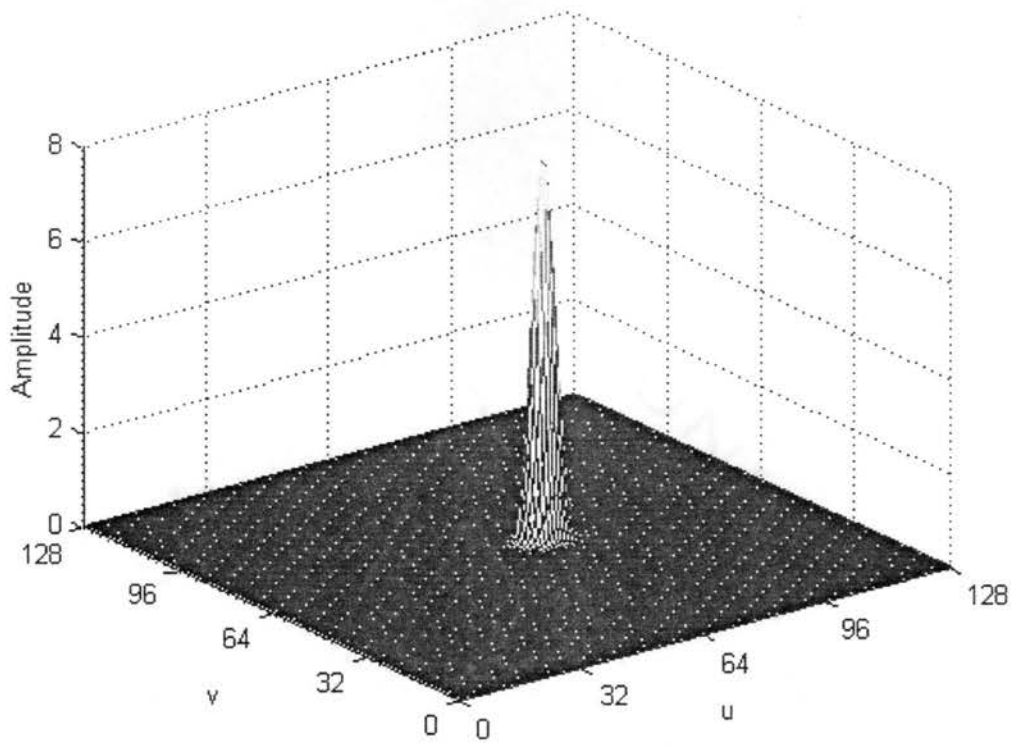


Figure 6.3. Frequency response of the Gabor filter ($K=3$; $S=3$; $n=0$; $m=0$).

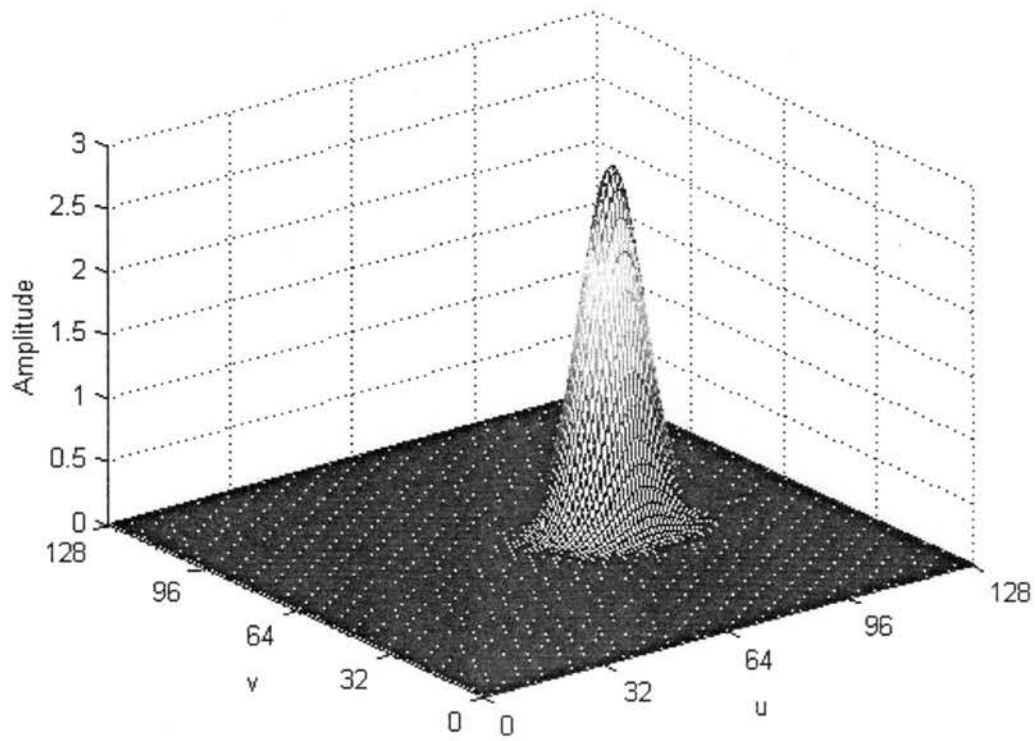


Figure 6.4. Frequency response of the Gabor filter ($K=3$; $S=3$; $n=0$; $m=1$).

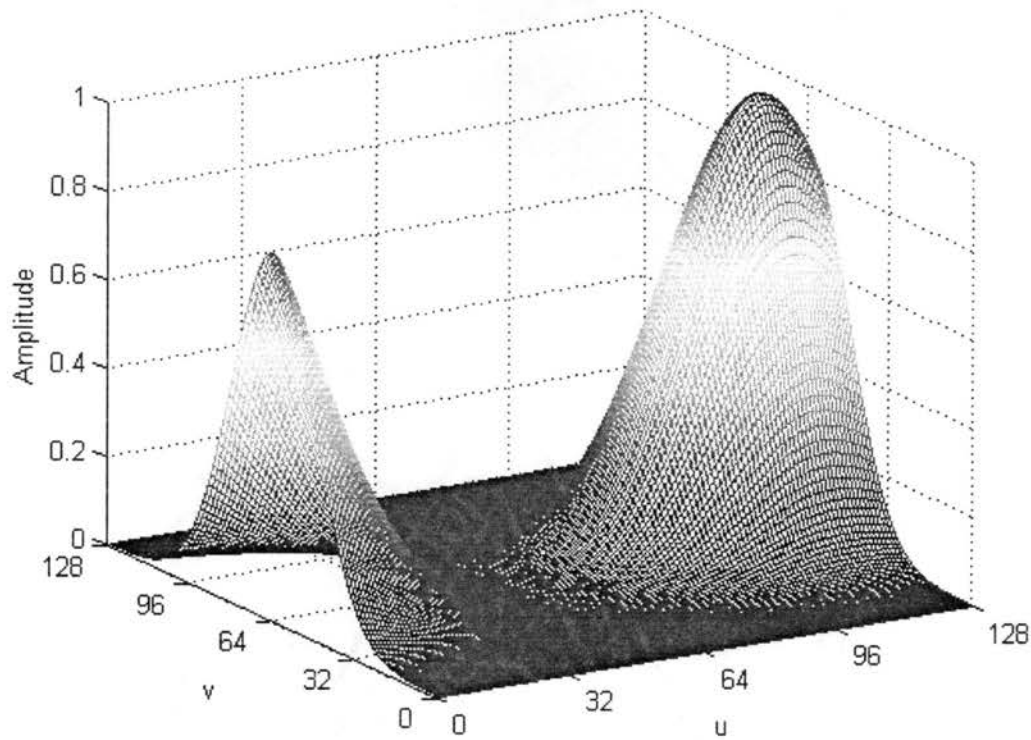


Figure 6.5. Frequency response of the Gabor filter ($K=3$; $S=3$; $n=0$; $m=2$).

As indicated earlier, frequency response of a filter defines what frequency regions are attenuated or amplified. Peak amplitude for coarse scale is higher than that for fine scale. The filters are designed such that the volume under the Gaussian for any scale and orientation is the same. Equal weight is given for all textural scales and orientations. The scale ($m = 0$) would capture coarse texture, whereas the scale ($m = 2$) would capture fine-textured features.

Following Figure 6.2, the Gabor filters for the other two orientations are given in Figures 6.6 and 6.7. All these filters would constitute a bank of band-pass filters. Gabor filters form a complete, but non-orthogonal basis set.

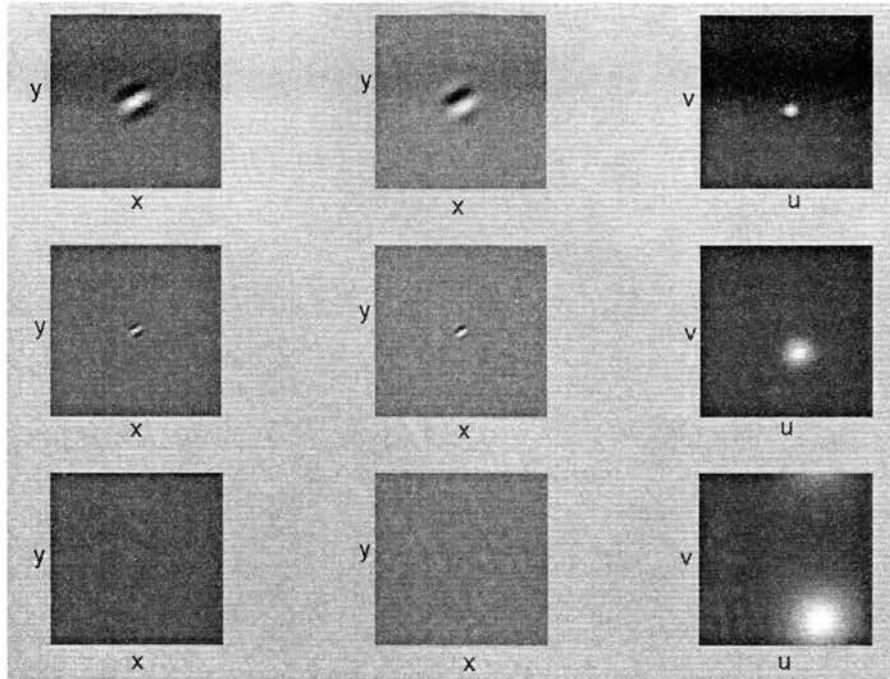


Figure 6.6. Impulse and frequency response of Gabor filter ($K=3$; $S=3$; $n=1$; $m=0, 1, 2$).

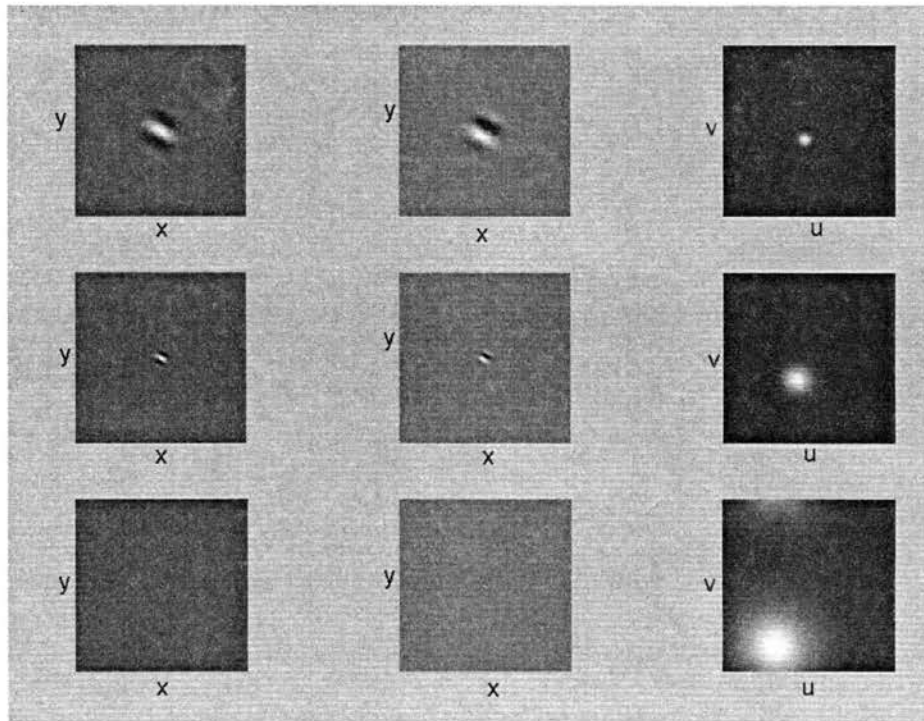


Figure 6.7. Impulse and frequency response of Gabor filter ($K=3$; $S=3$; $n=2$; $m=0, 1, 2$).

In this study, a Gabor filter with six orientations and four scales was used. Coverage of the frequency spectrum by a bank of band-pass filters is shown in Figure 6.8. Note that this figure is an image-display of frequency coverage, whereas Figure 6.1 displays half-peak magnitude contours of all band-pass filters. The volume under each Gaussian is the same, because the kernel function in Equation 6.1 contains the normalizing factor a^{-m} .

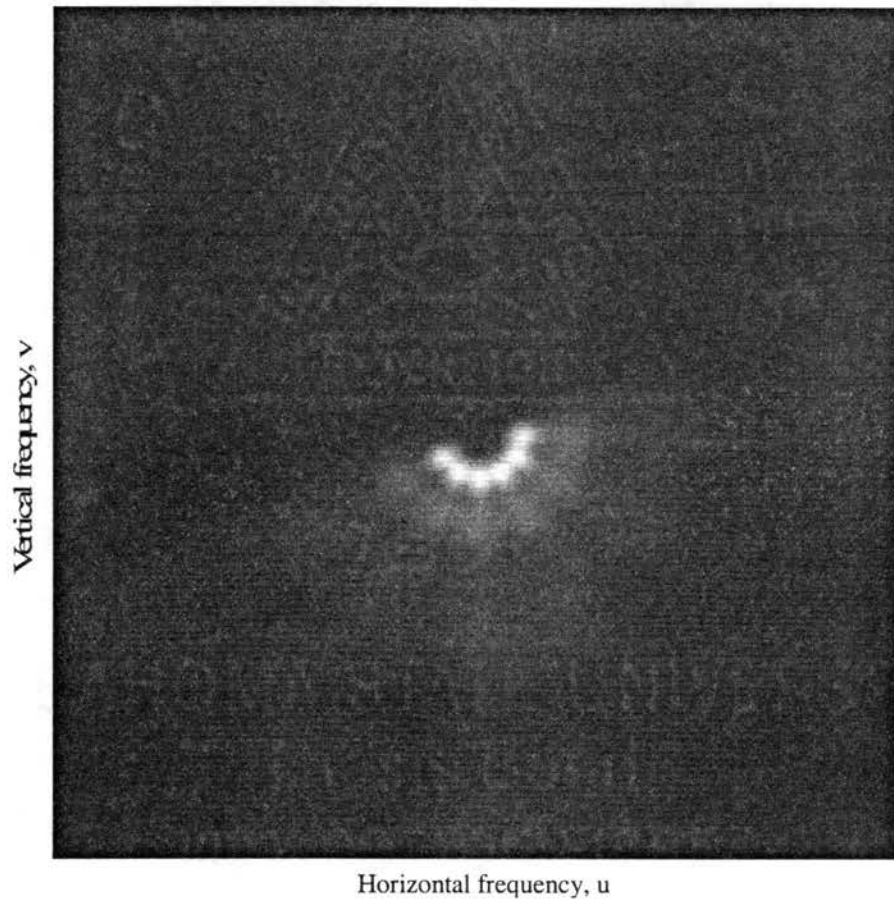


Figure 6.8. Coverage of frequency magnitude spectrum by Gabor filter.

Implementation of Gabor Filter

Implementation of a four-scale, six-orientation Gabor filter for a particular scale ($m = 1$) and horizontal orientation ($n = 0$) is illustrated in Figure. 6.9.

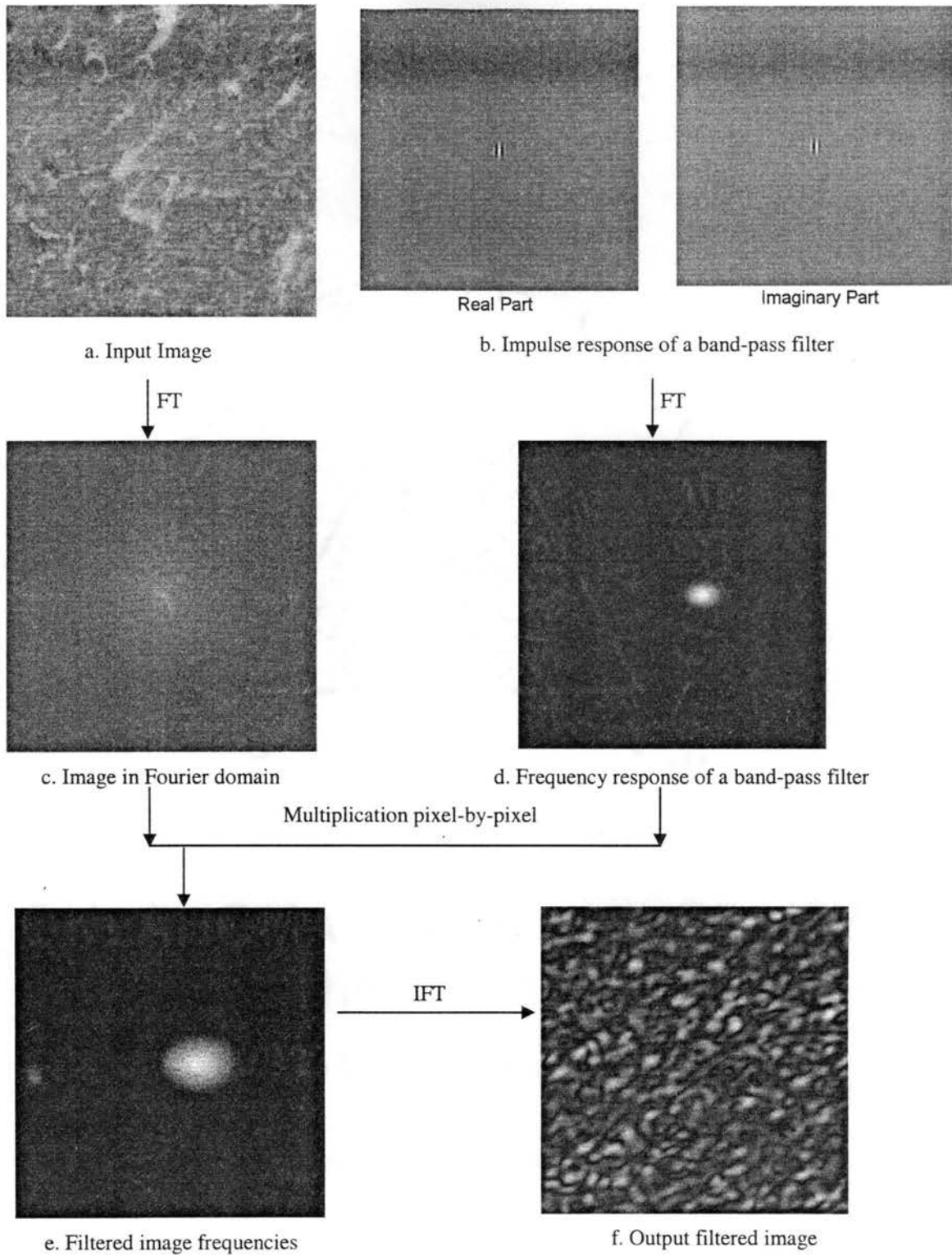


Figure 6.9. Illustration of a Gabor filter ($S=4$, $K=6$, $m=1$; $n=0$).

If the input image (Fig. 6.9a) is convolved with the impulse response of a band-pass filter (Fig. 6.9b) described by the Equation 6.2., the output image (Fig. 6.9f) will result. As indicated earlier, the convolution operation is computationally intensive. Filtering, therefore, is conducted in the Fourier domain. The FT identifies frequency components in the image (Fig. 6.9c). The FT of the impulse response reveals the frequency response of the filter. Frequency response clearly indicates the band of frequencies passed by the filter (Fig. 6.9d). When an FT of the image is multiplied pixel-by-pixel by the frequency response of the filter, the filtered image in the Fourier domain results (Fig. 6.9e). Note that the FT values have a large dynamic range. To display the FT as an image, the dynamic range is linearly compressed between 0 and 255. Dynamic ranges of Fig. 6.9d and 6.9e were different and were therefore compressed at different levels. Due to these differences in compression level, the frequency band in Fig. 6.9e appears to be larger than that in Fig. 6.9d. In fact, Figures 6.9d and 6.9e have the same frequency band. The filtered image in the Fourier domain (Fig. 6.9e) is then transformed by an IFT to the spatial domain (Fig. 6.9f).

If a particular location in the filtered image is bright, that location has more texture corresponding with that target frequency band and orientation. The mean of the filtered image represents the amount of texture described by the frequency band and orientation present in the input image. The variance of the filtered image would describe the variability of that texture from pixel-to-pixel in the image. The mean and the variance of the filtered images provide a description of textural distribution at various scales and orientations. In total, 48 features were extracted.

Canonical Discriminant Analysis

In this study, 186 beef samples were used and 48 features were extracted from the image of each sample. These beef samples were classified into three tenderness categories; tough, medium, and tender. Prediction of category would be simpler if all samples in each category formed a distinct cluster in the 48-dimensional feature space. This method would work successfully if there were large numbers of samples (10-20 times the number of features) and if features were uncorrelated. However, textural features are correlated, and a small number of samples can lead to an over-fitted model resulting in poor generalization.

Canonical discriminant analysis (CDA) is a dimension-reduction technique similar to principal component analysis (PCA). Like PCA, CDA also creates new uncorrelated features which are linear combinations of the original features. The number of new features generated is equal to the number of categories-1. For this study, two new canonical features were generated for each sample. Thus, CDA converted the original 48-dimensional feature space to a new 2-dimensional canonical feature space. The new features were calculated such that all samples within a category were tightly clustered and that the centers of the cluster categories were far apart. The new variables were then used in discriminant analysis (Johnson, 1998). The canonical discriminant model was developed in SAS (SAS Institute Inc., Cary, NC).

For classifying a new sample into a tenderness category, two canonical scores were calculated from 48 textural features using a linear combination equation. The new sample was then assigned to the category with the cluster center closest to the canonical scores of the new sample.

Results

Each image was decomposed to 24 images with distinct frequency bands. Calculating the mean and standard deviation of those 24 images produced 48 features. From the canonical model, two canonical features were calculated for all images. The mean canonical score for each category was then calculated as shown in Figure 6.10. We can divide the canonical feature space into three regions for three categories. The boundary lines are obtained by constructing perpendicular bisectors of lines connecting all pairs of group means and then erasing the portions of these bisectors that are not informative.

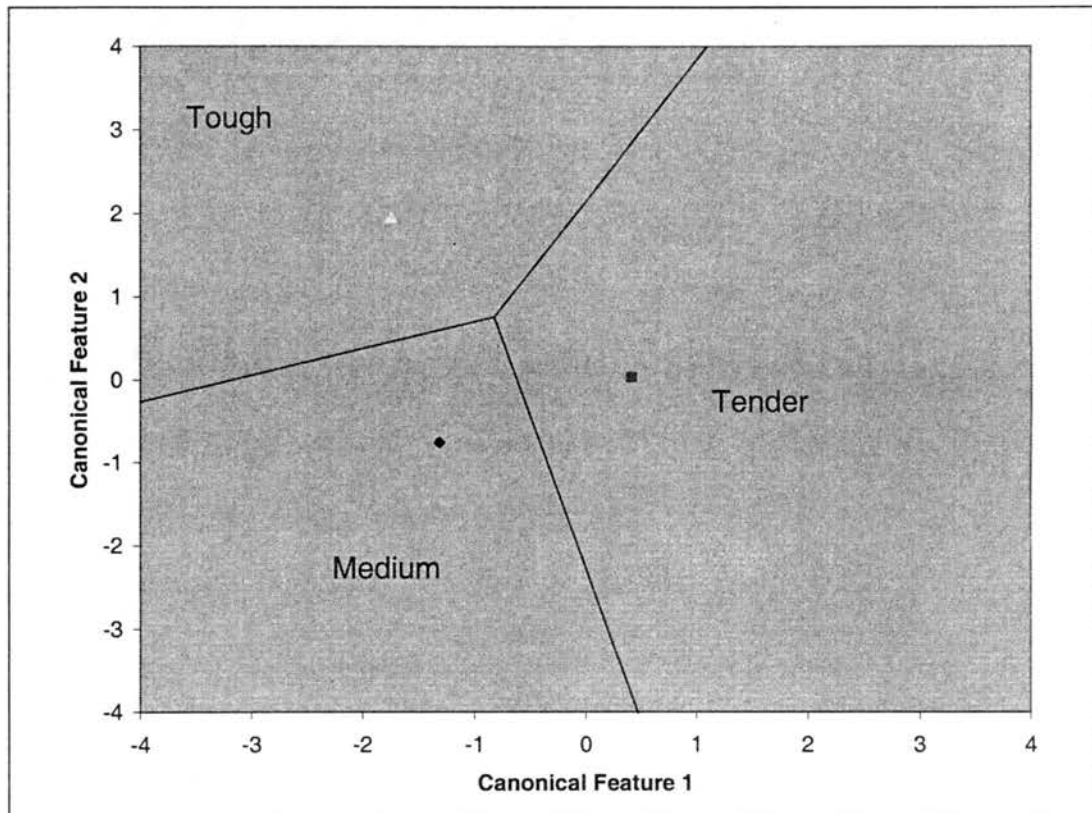


Figure 6.10. Territorial map for tenderness classification.

Figure 6.11 shows the canonical scores for three categories in the canonical feature space. As expected, there were a few misclassifications that fell on the category boundaries.

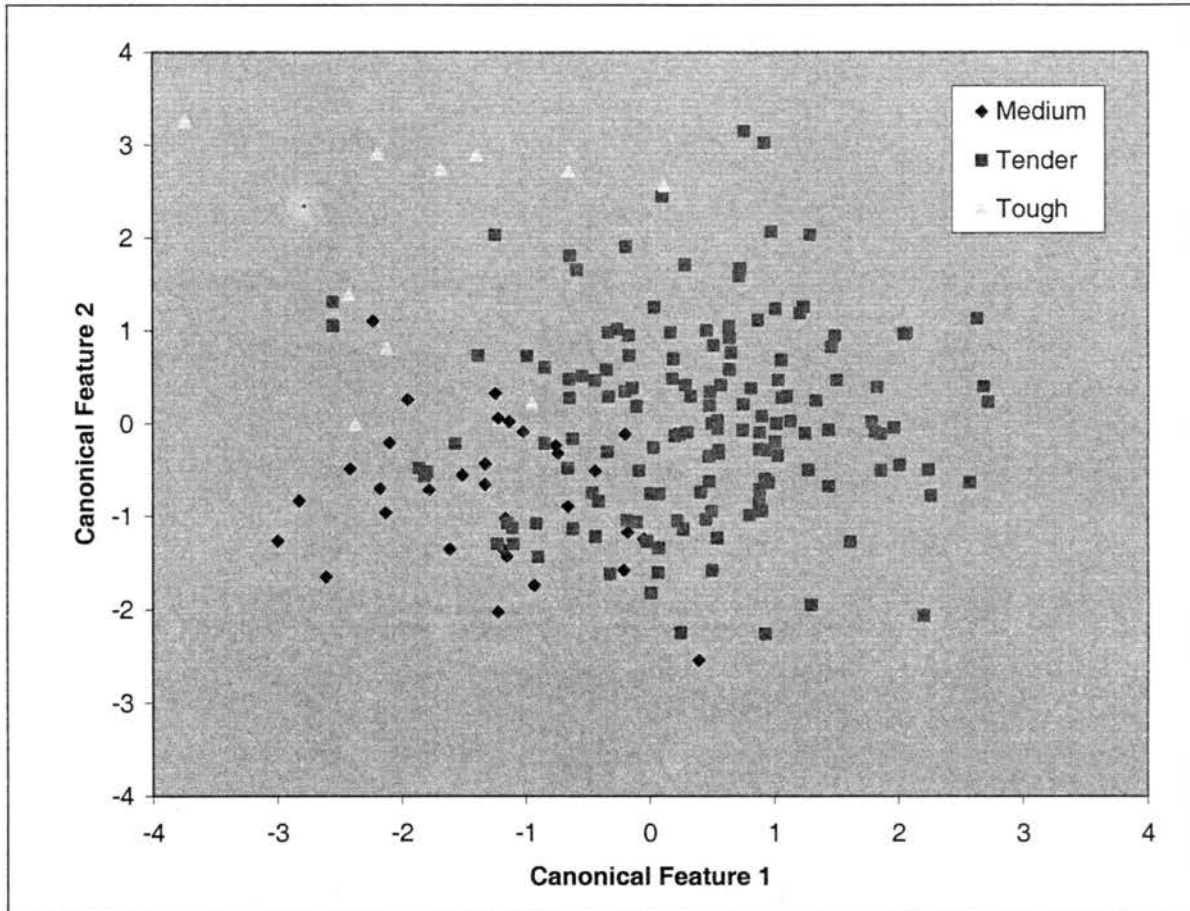


Figure 6.11. Distribution of beef samples in canonical feature space.

Table 6.1 shows the distance between the means of canonical features of three categories. As expected, “tender” is closer to “medium” than “tough.” However, the distance between the tough and tender is not much larger than the distance between tough and medium. This proximity could cause difficulty in discriminating new samples. Resubstitution results are obtained when the samples used for building the model are

employed to test the model (Table 6.2). Table 6.2 shows that 77.1% of tender steaks were correctly classified. Overall accuracy was 80.1%.

Table 6.1. Distance matrix between categories in canonical feature space.

Category	Tender	Medium	Tough
Tender	0	3.62	8.31
Medium	3.62	0	7.53
Tough	8.31	7.53	0

Table 6.2. Resubstitution results of canonical discriminant model.

Observed Category	Predicted Category			
	Tender	Medium	Tough	Total
Tender	111	22	11	144
	77.1%	15.3%	7.6%	100%
Medium	1	30	1	32
	3.1%	93.8%	3.1%	100%
Tough	0	2	8	10
	0%	20%	80%	100%
Accuracy	99.1%	55.6%	40%	80.1%

For testing, the cross-validation method was used. This procedure is also known as the “leave-one-out” method. The first sample was removed from the data set, and a discrimination rule was found using the remaining N-1 samples. The classification of the first sample was predicted. Then, the first sample was put back into the data set, and the second sample was removed. As before, a discrimination rule was found using the

remaining N-1 samples, and the classification of the second sample was predicted. This process was repeated for each sample. The resulting N different classifications can be used to find a nearly unbiased estimate of the accuracy of the discrimination rule (Johnson, 1998). Results are shown in Table 6.3.

Table 6.3. Cross-validation results.

Observed Category	Predicted Category			Total
	Tender	Medium	Tough	
Tender	111 77.1%	22 15.3%	11 7.6%	144 100%
Medium	2 6.3%	29 90.6%	1 3.1%	32 100%
Tough	0 0%	2 20%	8 80%	10 100%
Accuracy	98.2%	54.7%	40%	79.6%

Overall accuracy was 79.6%, closely approximating the 80.1% obtained for the resubstitution method. This match indicates that the model has not over-fitted and is therefore robust. The only concern with this model is that 11 tender samples were misclassified as tough. Over 98% of samples classified as tender were indeed tender samples. These samples can be sold as “guaranteed tender” to premium markets such as restaurants. Over 96% of samples classified as medium were either tender or medium. These samples can be directed to consumer markets. About 80% of tough carcasses were correctly identified in the tough category. About 10% of samples tested were predicted

in the tough category. These carcasses can be used for either processed meat or marketed after treating to improve tenderness.

Conclusion

We developed algorithms employing the Gabor filter to extract textural features that predicted aged-beef tenderness scores with 79.6% accuracy. Over 98% of samples classified as tender were indeed tender samples. About 80% of tough carcasses were correctly identified. The remaining 20% were misclassified in the neighboring tenderness category, medium. Implementation of our algorithms in existing computer vision systems for beef quality grading would enable the U.S. beef industry to expand market share by responding to consumer tenderness preference at the market place.

Acknowledgements

The authors thank the Food and Agriculture Products Research Initiative Program FRIP 2001, Oklahoma State University for supporting this study. We also thank Dr. Brad Morgan, OSU Animal Science, for providing his expertise in this project. Appreciation is expressed to Aaron Elam, Chad Carr, and Keith Charmasson, Animal Science Graduate Research Assistants for generating the shear-force data.

References

- AMSA. 1995. *Research guidelines for cookery, sensory evaluation, and instrumental tenderness measurements of fresh meat*. Savoy, IL.: National Live Stock & Meat Board, American Meat Science Association.
- Belk, K. E., J. D. Tatum, G. C. Smith, M. Goldberg, A. M. Wyle, and R. C. Cannell. 2001. Meat imaging system for palatability yield prediction. U.S. Patent No. 5793893.
- Benn, A., D. Barrett-Lennard, and P. J. Hay. 1998. Image analysis for meat. U.S. Patent No. 5793879.
- Boleman, S. L., S. J. Boleman, W. W. Morgan, D. S. Hale, D. B. Griffin, J. W. Savell, R. P. Ames, M. T. Smith, J. D. Tatum, T. G. Field, G. C. Smith, B. A. Gardner, J. B. Morgan, S. L. Northcutt, H. G. Dolezal, D. R. Gill, and F. K. Ray. 1998. National Beef Quality Audit-1995: survey of producer-related defects and carcass quality and quantity attributes. *Journal of Animal Science* 76(1): 96-103.
- Brooks, J. C., J. B. Belew, D. B. Griffin, B. L. Gwartney, D. S. Hale, W. R. Henning, D. D. Johnson, J. B. Morgan, F. C. Parrish, Jr, J. O. Reagan, and J. W. Savell. 2000. National Beef Tenderness Survey-1998. *Journal of Animal Science* 78(7): 1852-1860.
- Jeremiah, L. E. 1996. The influence of subcutaneous fat thickness and marbling on beef : Palatability and consumer acceptability. *Food Research International* 29(5-6): 513-520.

- Johnson, D. E. 1998. *Applied Multivariate Methods for Data Analysis*. New York, NY.: Duxbury Press.
- Lorenzen, C. L., D. S. Hale, D. B. Griffin, J. W. Savell, K. E. Belk, T. L. Frederick, M. F. Miller, T. H. Montgomery, and G. C. Smith. 1993. National Beef Quality Audit: survey of producer-related defects and carcass quality and quantity attributes. *Journal of Animal Science* 71(6): 1495-1502.
- Manjunath, B. S., and W. Y. Ma. 1996. Texture features for browsing and retrieval of image data. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18(8): 837-842.
- Mckenna, D. R., D. L. Roebert, P. K. Bates, T. B. Schmidt, D. S. Hale, D. B. Griffin, J. W. Savell, J. C. Brooks, J. B. Morgan, T. H. Montgomery, K. E. Belk, and G. C. Smith. 2002. National Beef Quality Audit-2000: survey of targeted cattle and carcass characteristics related to quality, quantity, and value of fed steers and heifers. *Journal of Animal Science* 80(5): 1212-1222.
- Misiti, M., Y. Misiti, G. Oppenheim, and J. Poggi. 1996. Wavelet Toolbox for use with Matlab - User's Guide. Ver. 1. Natick, MA.: The Mathworks, Inc.
- NCBA. 2003. Economics of Beef. Beef Bytes. Centennial, CO.: National Cattlemen's Beef Association. Available at:
<http://www.beef.org/documents//Beef%20Bytes%20Economics.pdf>. Accessed on 24 July 2003.
- OBIC. 2003. Oklahoma Beef Statistics. Oklahoma City, OK.: Oklahoma Beef Council. Available at: <http://www.oklabeef.org/Beef%20Statistics.htm>. Accessed on 24 July 2003.

Oppenheim, A. V., and R. W. Schaffer. 1999. *Discrete-Time Signal Processing*. 2nd ed.
New Jersey, NJ.: Prentice Hall.

Tong, A. K., D. J. Robinson, and T. Liu. 1998. Method and apparatus for using image
analysis to determine meat and carcass characteristics. Canadian Patent No.
02263763.

USDA. 1997. United States standards for grades of carcass beef. Effective date January
31, 1997. Washington, DC.: Livestock and Seed Division of the Agricultural
Marketing Service. Available at:
<http://www.ams.usda.gov/lsg/stand/standards/beef-car.pdf>. Accessed on 9
February 2004.

USDA. 2001. Procedures for approval and use of instrument augmentation systems for
beef carcass ribeye measurement. Washington, DC.: Livestock and Seed Program,
Agricultural Marketing Service.

CHAPTER VII
PREDICTING BEEF TENDERNESS FROM
WAVELET TEXTURAL FEATURES

S. Jeyamkondan and G.A. Kranzler

Abstract

Beef tenderness is a primary consideration in consumer satisfaction. A computer vision system was developed to predict 14-day aged cooked-beef tenderness. Images of 186 fresh beef steaks were acquired at 3-d postmortem. Images were decomposed up to five levels using the fourth-order Daubechies wavelet. Energy and color correlation signatures were extracted from wavelet-decomposed images and used as textural features. A stepwise linear regression model was used to predict 14-d Warner-Bratzler shear-force (WBS) tenderness scores. The correlation coefficient between the observed and predicted shear-force values was 0.54. In comparison, 7-d WBS values predicted 14-d WBS values with a correlation coefficient of 0.53. Thus, predictions by textural features extracted from fresh images acquired at 3-d postmortem using the computer vision system (nondestructive method) were as accurate as 7-d WBS values (sample-destructive method).

The beef industry is more interested in identifying “tough” carcasses than in predicting shear-force values. A limited number of tough samples (10) in the tested

carcasses prevented evaluation of the system for classifying carcasses into “tender” and “tough” categories. An alternative method was used to evaluate the system for sorting carcasses based on tenderness at nine certification levels. At all tenderness certification levels, there were significant differences between the shear-force values for “certified tender” and “not certified tender” samples. The system was capable of accurately sorting and categorizing fresh beef samples on the basis of aged-beef tenderness.

Keywords. Daubechies wavelet, short-time Fourier transform, computer vision, quadrature mirror filters, Warner-Bratzler shear force, image processing, video image analysis.

Introduction

In 1995, the National Cattlemen’s Beef Association listed, “..development of an instrument or procedure that can adequately measure quality, cutability and tenderness in beef carcasses in modern packing plants...” as a top priority of the beef industry. During the last decade, extensive research has been directed toward developing instruments for measuring beef palatability. Significant advances were made in predicting cutability, but not in estimating tenderness.

Tenderness is a critical factor in consumer perception of beef palatability. In the 1995 National Beef Quality Audit (Boleman et al., 1998), inadequate tenderness was identified as the second largest quality concern in the beef industry. The National Beef Tenderness Survey (Morgan et al., 1991) showed that consumers experience undesirable toughness in 25% of steaks consumed. Despite its significance as a quality indicator, tenderness is not a factor affecting product value for producers and packers, because the USDA quality grading system (USDA, 1997) does not currently incorporate tenderness.

Direct evaluation is absent, because there is currently no widely accepted method available for predicting tenderness online. Because carcasses are not priced on the basis of tenderness, producers lack incentive to supply a tender product. As a result, consumer preference is not routed back to the producers. There is growing recognition that beef tenderness must be incorporated in the quality grading process if true, value-based marketing is to develop.

Responding to the need for objective measurement, researchers have developed three instrumented systems for predicting beef tenderness: BeefCam[®] (Belk et al., 2001), Colorimeter Method (Wulf et al., 1997), and Slice Shear Force (Shackelford, et al., 1997). In a recent comparative study, performance of these systems was evaluated. Because of low accuracy or invasive requirements, none of the systems was judged ready for industry acceptance (Wheeler et al., 2002). A sufficiently accurate, nondestructive method for online evaluation of tenderness continues to elude the beef industry.

Kim et al. (1998) utilized wavelet textural features from ultrasound images to predict intermuscular fat percentage with an R^2 value of 0.62. Huang et al. (1997) extracted 25 wavelet textural features for beef tenderness prediction from 29 ultrasonic elastogram images. The wavelet textural features predicted WBS scores obtained after aging for 2, 14, 28, and 42 days with R^2 values ranging from 0.72-0.95. Had 26 samples been used instead of 29, they would have obtained a perfect R^2 value of 1. The large number of parameters in the regression model, with respect to number of samples tested, raised serious questions of overfitting.

The USDA has recently approved video image analysis (VIA) or computer vision for predicting yield grade in packing plants (USDA, 2001). Commercial VIA systems

are available to predict yield grades (Benn et al., 1988; Tong et al., 1998; Belk et al., 2001). These systems capture and process ribeye images to predict yield and quality grades. The opportunity exists to extract textural features from those ribeye images to predict beef tenderness.

Li et al. (2001) used wavelet-based textural features to classify beef carcasses into tender and tough groups. Fifty-nine samples were imaged with a spatial resolution of 480 x 512. Each image was divided into 56 subimages, with 64 x 64 resolution. From eight extremely tender samples, 45 subimages were randomly selected. Another 45 subimages were randomly selected from nine extremely tough samples. These 90 subimages were the only images used in the study (Li et al., 1999). It is not clear why the authors did not use more subimages in this study, when one tough sample can produce 56 subimages. Wavelet textural features (n=35) were extracted from each subimage. Fisher's discriminant model was developed to classify tender and tough samples and was evaluated by cross-validation (leave-one-out technique). They achieved 83.3% correct classification (Li et al., 2001; Li et al., 1999). Only 17 samples with extreme tenderness scores were used in this study. Had the other samples with intermediate tenderness scores been used, the correct classification rate would obviously have been reduced.

In summary, studies by Li et al. (2001) and Huang et al. (1997) show promise for utilizing wavelet textural features for beef tenderness prediction. The objective of this study was to develop a computer vision system that can extract textural features in the wavelet domain from fresh beef images to predict aged, cooked-beef tenderness.

Materials and Methods

Samples

Beef ribeye rolls (n=186) were collected from regional packing plants.

Descriptive statistics from sample carcass data are shown in Table 7.1.

Table 7.1. Descriptive statistics from sample carcass data.

Parameter	Mean	Standard Deviation	Minimum	Maximum
Marbling ^a	494.8	72.9	330	730
Quality Grade ^b	18.0	1.4	15	21
7-d WBS ^c , kg	3.80	0.79	2.50	7.16
14-d WBS, kg	3.41	0.64	2.25	5.81

^aMarbling: 300=Traces⁰⁰, 400=Slight⁰⁰, 500=Small⁰⁰, 600=Modest⁰⁰, 700=Moderate⁰⁰

^bUSDA Quality Grade: High Standard=15, Low Select=16, Average Select=17, High Select=18, Low Choice=19, Average Choice=20, High Choice=21 (Johnson and Dockerty, 1990).

^cWBS: Warner-Bratzler shear-force tenderness values.

Marbling scores ranged from Traces to Moderate. Quality grade ranged from High Standard to High Choice. Most of the carcasses graded either Select or Choice, which is in agreement with grade distribution identified in the most recent National Beef Quality Audit – 2000 (McKenna et al., 2002). As expected, the mean and standard deviation of the Warner-Bratzler shear-force (WBS) values on 14-d postmortem samples were smaller than those for 7-d postmortem. These results indicate that longer-aged samples were more tender, and that variability of tenderness among samples was reduced with aging.

Two one-inch paired steaks were cut from each ribeye roll and individually identified. The steaks were allowed color-bloom for 30 minutes. One of the steaks was

imaged, then vacuum packaged and aged for 14 days. The other steak was vacuum packaged and aged for 7 days.

Computer Vision System

The computer vision system consisted of a color video camera (A209 MicroImage Video Systems), an image digitizer (Flash Point 128 Integral Technologies) and a 1.0 GHz PC and monitor. Images of the beef steaks were captured in a vaulted lighting chamber. Diffuse, uniform light was provided by six 50-watt halogen lamps powered by a stabilizing feedback controller. The camera was mounted above the chamber, viewing a 32 x 43 mm imaging area through an observation port. Camera resolution was 640 x 480 pixels. Pixel size was 67 x 67 μm . Output was in the red-green-blue (RGB) format.

Tenderness Reference Measurement

The imaged steak was designated for obtaining 14-day aged WBS values, whereas 7-day aged WBS values were collected from the other steak from the same carcass. Following aging at 1° C, steaks were cooked in an impingement oven to an internal temperature of 70° C. After cooling, six 12-mm diameter cores were sampled randomly from each steak ribeye. Force required to shear the core was recorded using a Warner-Bratzler attachment to an Instron Universal Testing Machine. Peak shear-force values from the six replicate cores were averaged to obtain the shear-force tenderness reference value (AMSA, 1995). Higher shear-force values indicated “tougher” beef.

Image Texture

References on texture analysis often begin with the observation that there is no single, unambiguous, widely accepted definition for texture (Livens, 2002). IEEE

standard 610-4 (1990) states, “Texture is an attribute representing the spatial arrangement of the gray levels of the pixels in a region.” Texture refers to repetition of basic patterns called texels (Jain, 1989). Textures in nature are often random. Because there is no clear definition for texture, extracting textural features has been an ongoing challenge. In this study, textural features were extracted in the wavelet domain. Image processing algorithms were coded in Matlab 6.0 (Mathworks, Natick, MA).

Need for New Transform

The spatial domain (original image) contains complete spatial (temporal, in case of signal processing) information, but no frequency information. The Fourier transform (FT) identifies all spectral components of an image; however it does not contain spatial information. The FT analyzes the image/signal globally, not locally. In order to obtain spatial localization of frequency components, the image must be analyzed locally. To meet this requirement, the short-time Fourier transform (STFT) was introduced (Oppenheim and Schaffer, 1999).

Short-time Fourier Transform

To begin implementation of the STFT, a window of finite length is selected. This window is moved over the signal, and the FT of the windowed signal is calculated. The STFT is a compromise between temporal and frequency domains. When the window size is infinitesimally small, the STFT becomes a unit impulse function and gives only temporal information. When the window is infinitely large, the STFT becomes an FT. The wider the window, the lower the temporal resolution and the higher the frequency resolution. The Heisenberg uncertainty principle implies that it is not possible to achieve

higher temporal resolution and higher frequency resolution at the same time. The main disadvantage of the STFT is that once the window size is fixed, spatial/temporal resolution is the same for all frequency components (Pollikar, 2002).

Wavelet Transform

The higher the frequency, the higher is the rate of change in gray level with respect to spatial distance. In order to capture high frequencies, high spatial resolution is required. Typically, an image contains very few high frequencies, and thus larger bandwidth (low-frequency resolution) is adequate. Low frequencies indicate that the gray level changes slowly, and therefore low spatial resolution is sufficient. A high-frequency resolution (narrow bandwidth) is required, because an image contains a large amount of low-frequency content. It follows that the resolutions in the frequency and spatial domains are inversely related.

It is preferred to have a transform with variable window size, instead of a fixed window size as in the STFT. The wavelet transform (WT) overcomes the limitations of the STFT by employing a variable-length window. A narrower window is employed in the high frequency range, because it provides higher spatial resolution (lower frequency resolution) to capture high-frequency content of the image. A wider window is used in the low-frequency range, because it provides higher frequency resolution (lower spatial resolution) to analyze low-frequency content of the image. The function used to window the image is called a “mother wavelet.”

The WT research community uses the term “scale” instead of “frequency.” Scale is defined as $1/\text{frequency}$. As with maps, high-scale provides an overview of the region

(low-frequency content), whereas low-scale provides intricate details of the region (high-frequency content).

Continuous Wavelet Transform

A one-dimensional continuous wavelet transform (CWT) for a time signal is defined as (Van de Wouwer, 1999):

$$W(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \cdot \psi^* \left(\frac{t-b}{a} \right) dt \quad (7.1)$$

Where: $W(a,b)$ = CWT coefficients of a time signal, $x(t)$,

a = translation parameter, measure of time,

b = scale parameter, inversely related to frequency.

$$\psi^*_{a,b}(t) = \psi^* \left(\frac{t-a}{b} \right) = \text{mother wavelet } \psi^* \text{ translated by 'a' and scaled by 'b'}$$

Translation means a time-shift of the signal. When the scale is between 0 and 1 (low-scale), the mother wavelet ψ^* is contracted (narrow window) and can analyze high frequency. When the scale is greater than 1 (high-scale), the mother wavelet is expanded or dilated (wider window) and can analyze low frequency. In many practical applications, high frequency does not continue for the duration of the signal. Thus, we need high temporal resolution to capture the high-frequency components (fast-changing signal). Low frequency usually continues for the duration of the signal. We need less temporal resolution to capture slowly-changing signal components (Pollikar, 2002).

CWT can be interpreted as a time-scale plot of the signal. The mother wavelet is translated by 'b' and dilated by 'a' and then multiplied by the entire signal. The result is normalized by dividing by the square root of the scale, so that the transformed signal has equal energy at all scales. If the signal at the particular location (b) has a frequency

component corresponding to that scale (a), the CWT coefficient, $W(a,b)$, will be large. Therefore, $W(a,b)$ can be interpreted as the cross-correlation between the signal and the mother wavelet translated by 'b' and scaled by 'a' (Pollikar, 2002).

Mother Wavelet

The term 'wavelet' literally means small wave, indicating that it is of finite width and therefore compactly supported. The wavelet functions should satisfy the following conditions (Pollikar, 2002):

$$\int_{-\infty}^{\infty} \psi(t) dt = 0 \quad (7.2)$$

$$\int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty \quad (7.3)$$

Equation (7.2) signifies that it is a wave with a zero mean. It does not have a D.C. component. Equation (7.3) signifies that the energy of the wavelet function is finite, meaning that it is compactly supported. Note that in the FT, the base functions are sines and cosines, which have infinite width. An example of a mother wavelet is shown in Figure. 7.1.

The Haar wavelet is the earliest and simplest wavelet. It is a square step function. Other wavelet families include Daubechies, biorthogonal, coiflets, symlets, Morlet, Mexican hat, and Meyer (Misiti et al., 1996). In short, wavelets are mathematical functions that decompose the signal or image into various frequency components with resolution matched to the scale.

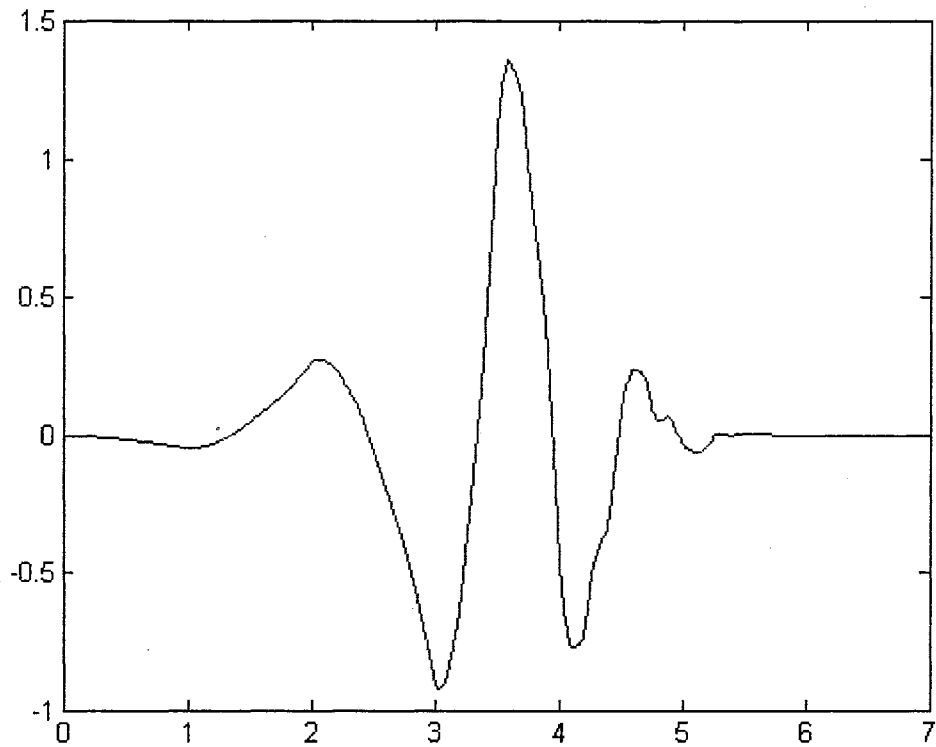


Figure 7.1. Fourth-order Daubechies (db4) mother wavelet.

Discrete Wavelet Transform

Even though the CWT can be implemented in discrete format, there is too much redundant information in the decomposed signals/images. In the discrete wavelet transform (DWT), down-sampling is performed at each level of decomposition to remove redundant information. At the same time, perfect reconstruction of the original signal/image is possible.

Filter-bank Representation of DWT

Discrete wavelet transforms are usually implemented using filters. $W(a,b)$ can be interpreted as the output of the filter with an impulse response of $\psi_{a,0}^*(-b)$ to the input signal $x(b)$. Thus, there is a family of filters characterized by the translation parameter 'a' for a given mother wavelet. For more details, refer to Strang and Nguyen (1996).

Let $x(n)$ be the original signal, sampled at n points. Let $h(n)$ and $g(n)$ be low- and high-pass filters corresponding to a given mother wavelet. Low- and high-pass filters for the Daubechies of order 4 (db4) wavelet are given in Figures 7.2. In fact, the low-pass filter ($h[n]$) and high-pass filter ($g[n]$) are related to each other by the following relationship (Pollikar et al, 2002):

$$h[L-1-n] = (-1)^n g[n], \quad (7.4)$$

where L is the filter length. This relationship allows for perfect reconstruction of the original signals from wavelet coefficients. The filters satisfying the above condition are known as “quadrature mirror filters.”

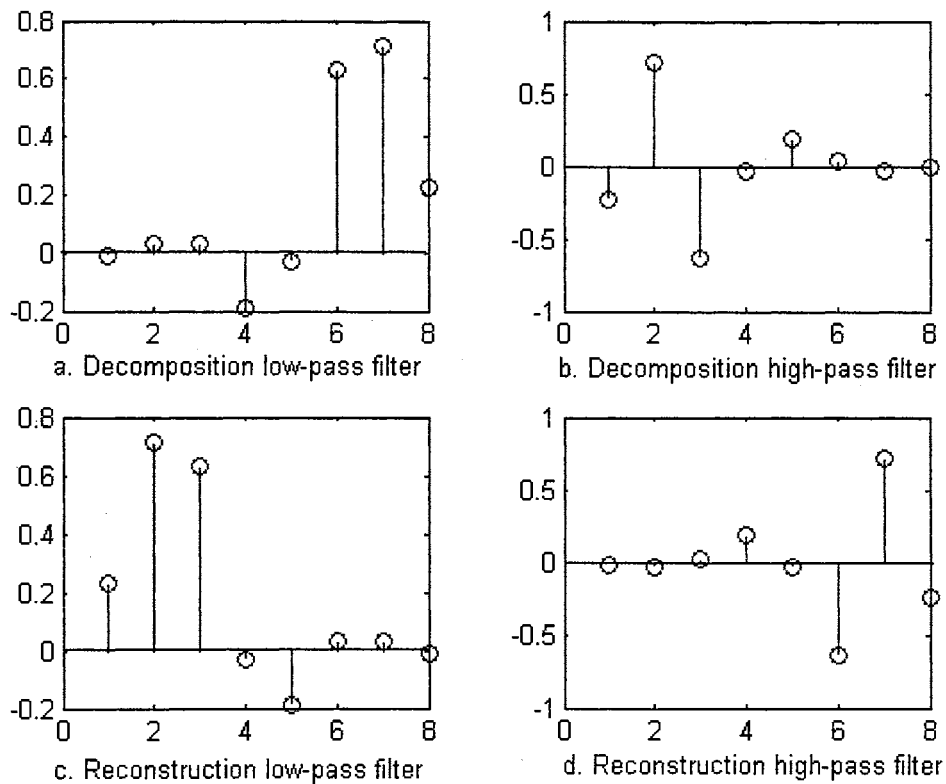


Figure 7.2. Low- and high-pass filters for implementation of the 4th order Daubechies wavelet.

The original signal is first passed through the low-pass filter to obtain approximation and the high-pass filter to obtain details. The resulting signals are down-sampled by a factor of 2. This procedure results in the same memory storage requirements for the decomposed signals as for the original signal. The approximation at level-1 decomposition can be sent through the same low- and high-pass filters to obtain the second-level DWT coefficients. The decomposition can be continued.

As the levels of decomposition increase, the amount of down-sampling increases. The result is lower resolution. At the first level of decomposition, high-frequency components are extracted with higher temporal/spatial resolution. As the number of levels increases, the lower-frequency components are extracted at lower temporal/spatial resolution. This sequence results in the decomposition of the original signal into a series of signals with different frequency components and with resolution matched to scale. This is known as “multiresolution analysis” (MRA).

Original signals can be perfectly reconstructed from the wavelet coefficients. Reconstruction low-pass ($\bar{h}[n]$) and high-pass ($\bar{g}[n]$) filters are related to decomposition low- and high-pass filters by (Figure 7.2):

$$\begin{aligned}\bar{h}[n] &= h[L-1-n] \\ \bar{g}[n] &= g[L-1-n]\end{aligned}\tag{7.5}$$

For reconstruction of the signals from wavelet coefficients, details and approximation coefficients at the last level of decomposition are first up-sampled by a factor of 2 and passed through the reconstruction low- and high-pass filters. Outputs from the two filters are added to obtain the approximation at the previous level. This sequence is continued until the original signal is obtained.

Wavelet Decomposition of Images

A two-dimensional WT can be obtained from the product of two one-dimensional WT's.

The one-dimensional WT provides two sets of coefficients at each level of

decomposition; namely, approximation and details. Likewise, a two-dimensional WT

provides 4 sets of coefficients at each level of decomposition (Fig. 7.3).

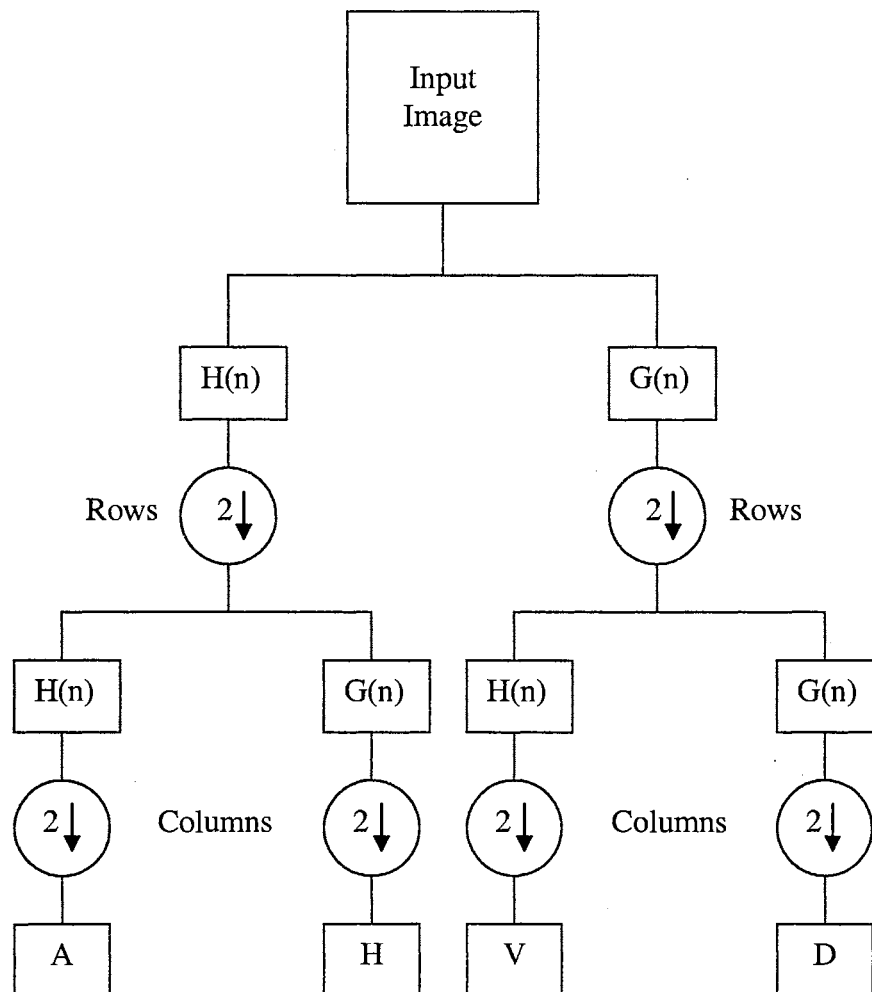


Figure 7.3. Wavelet decomposition of images using filters. Downward arrows indicate down-sampling by a factor of 2 along rows or columns. A - approximation, H - horizontal details, V - vertical details, D - diagonal details.

Rows of the input image are first passed through the low- and high-pass filters, followed by down-sampling along the rows by a factor of 2. The columns of resulting images

from both filters are sent through low- and high-pass filters followed by down-sampling along the columns to obtain four sets of coefficients; namely, approximation, horizontal, vertical, and diagonal details (Fig. 7.3).

The original image has now been decomposed using the db4 wavelet for level 1 (Fig. 7.4). For the next level of decomposition, the approximation (A1) is again decomposed to second-level approximation (A2), horizontal (H2), vertical (V2), and diagonal details (D2). This procedure can be continued for further decomposition.

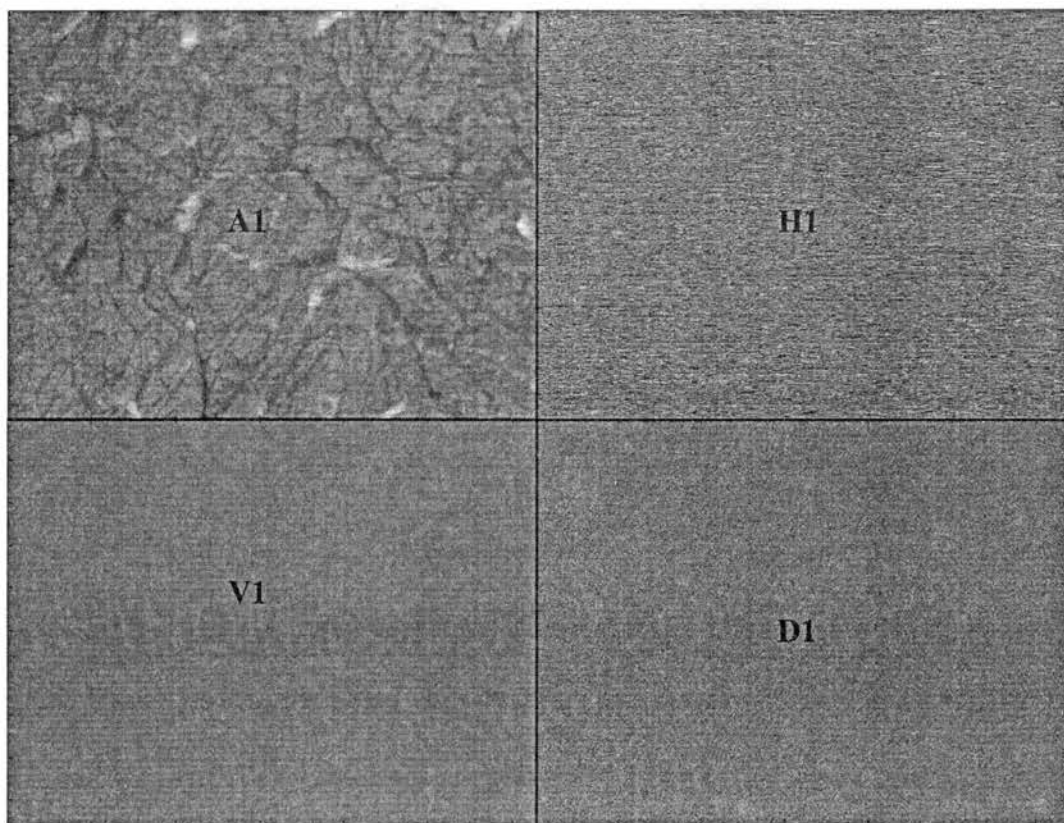


Figure 7.4. Level 1 decomposition of image using the Haar wavelet. A - approximation, H - horizontal details, V - vertical details, D - diagonal details.

Another way to view the wavelet decomposed images is to reconstruct the image using particular wavelet coefficients at each level. This approach produces the same-sized image for each approximation and detail. However, the high-level decomposed

images will show pixelization, because they were constructed from low-resolution coefficients (Fig. 7.5). Details at level 1 contain the highest frequencies of the original image, and therefore need higher spatial resolution to show the quickly varying details. As the level increases, the frequency of the details decreases. For example, level 5 has the coarsest or lowest frequency content of the image. Thus, lower resolution is sufficient to capture those slowly changing details. The original and decomposed images have the following relationship. Let I be the original image (Fig. 7.5).

$$\begin{aligned}
 I &= A_1+H_1+V_1+D_1 \\
 &= A_2+H_2+V_2+D_2+H_1+V_1+D_1 \\
 &= A_3+H_3+V_3+D_3+H_2+V_2+D_2 +H_1+V_1+D_1 \\
 &= A_4+ H_4+V_4+D_4+H_3+V_3+D_3+H_2+V_2+D_2 +H_1+V_1+D_1 \\
 &= A_5+ H_5+V_5+D_5+H_4+V_4+D_4+H_3+V_3+D_3 + H_2+V_2+D_2 +H_1+V_1+D_1
 \end{aligned}$$

Wavelet Features

One of the simplest and most effective features is the energy in wavelet-decomposed images. The (normalized) energy of a decomposed image of size $N \times M$ can be calculated by:

$$E = \frac{1}{NM} \sum_{i=1}^N \sum_{j=1}^M (I(i, j))^2 \quad (7.5)$$

Because details have a zero mean, the energy is also equal to the variance of the decomposed images. Wavelet energy signatures indicate the distribution of energy along the frequency axis over scale and orientation. They are powerful for characterizing image texture (Van de Wouwer et al., 1999). Energy signatures for color images can be obtained separately for each RGB color band.

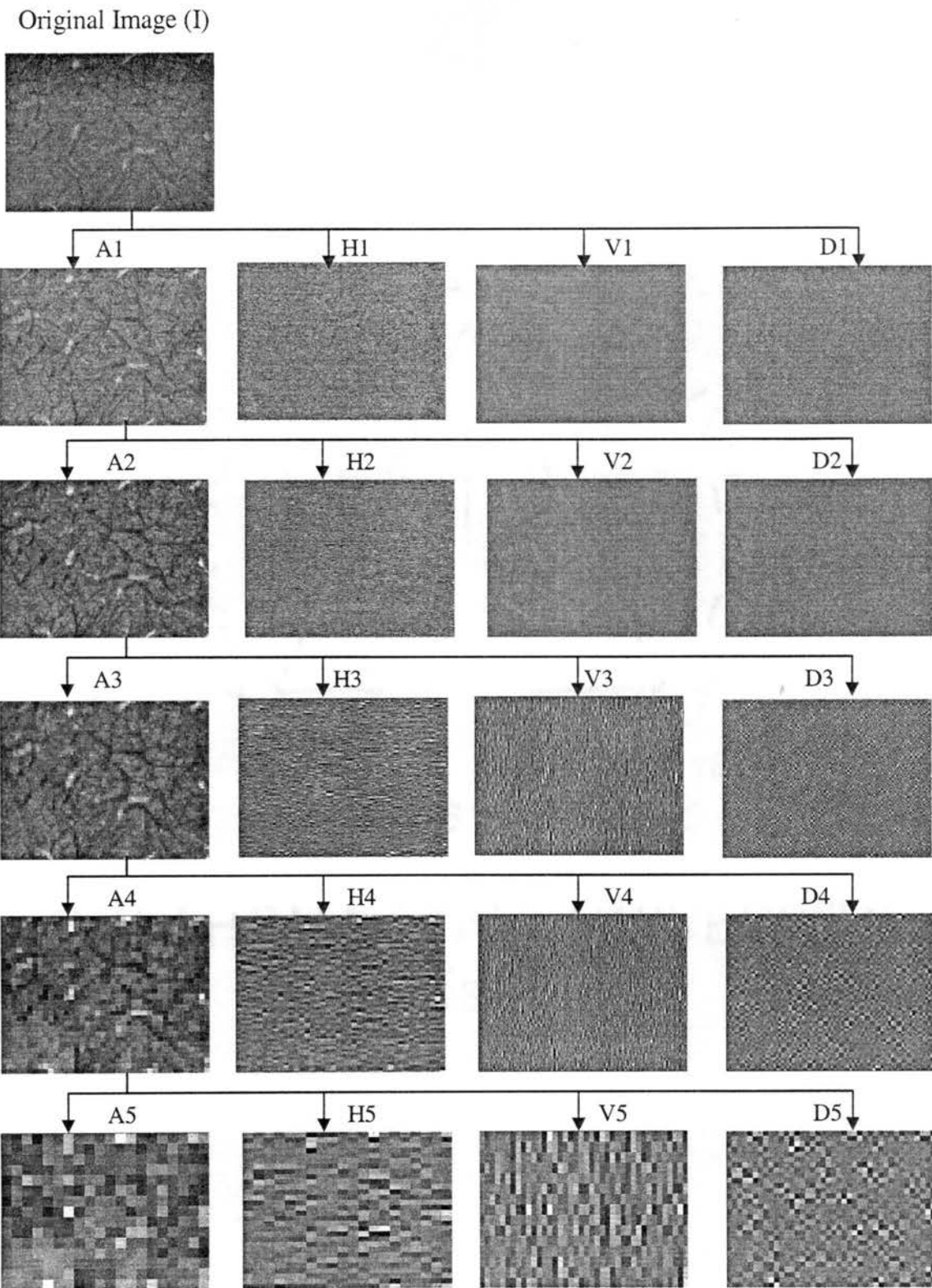


Figure 7.5. Five-level decomposition of beef images using the db4 wavelet.

Color covariance features can be calculated from the energy of the image that is obtained by multiplying the corresponding detail coefficients of two color bands pixel-by-pixel. Covariance features for wavelet-decomposed color image at a particular scale and orientation can be calculated by (Van de Wouwer et al., 1999):

$$\begin{aligned}
 C_{RG} &= \frac{1}{NM} \sum_{i=1}^N \sum_{j=1}^M (R(i, j)G(i, j)) \\
 C_{GB} &= \frac{1}{NM} \sum_{i=1}^N \sum_{j=1}^M (G(i, j)B(i, j)) \\
 C_{BR} &= \frac{1}{NM} \sum_{i=1}^N \sum_{j=1}^M (B(i, j)R(i, j))
 \end{aligned} \tag{7.6}$$

Where: C_{RG} = covariance feature between the red (R) and green (G) bands,

C_{GB} = covariance feature between the G and blue (B) bands,

C_{BR} = covariance feature between the B and R bands of the wavelet-decomposed image at a particular scale and orientation.

Covariance features are dependent on the energy of the corresponding color bands. Covariance features can be normalized by dividing by the square root of corresponding color-band energies to produce color correlation features as given by (Livens et al., 1997):

$$\bar{C}_{pq} = \frac{C_{pq}}{\sqrt{C_{pp}} \sqrt{C_{qq}}} \tag{7.7}$$

Where: $p = q = \{R, G, B\}; p \neq q,$

\bar{C}_{pq} = color correlation feature between the color bands, p and q,

C_{pq} = color covariance feature between the color bands, p and q,

C_{pp} = energy in the color band, p, of the wavelet-decomposed image.

Color correlation values range from -1 to 1. If the detail coefficients of two color-bands vary spatially in a similar fashion, the color correlation value will be close to 1.

Color Space Conversion

Color influences the perception of texture. In this study, wavelet textural features were extracted in HSI (hue, saturation, intensity) and CIE L*a*b* (LAB) color spaces in addition to RGB format. Intensity indicates the brightness of the image, which is calculated by averaging the R, G, and B color bands. Intensity ranges from black to white. Hue describes a pure color (yellow, orange, or red), whereas saturation describes the degree to which a pure color is diluted by white light. Details for converting RGB to HSI are outlined in Gonzalez and Woods (1992).

The CIE LAB is a perceptually uniform color space which mimics the logarithmic response of the eye. The lightness, L* ranges from 0 (black) to 100 (white), a* ranges from -100 (green) to 100 (red), and b* ranges from -100 (blue) to 100 (yellow). Details for conversion of RGB to CIE LAB color space are given in Connolly and Fleiss (1997). For conversion, chromaticity values from a white reference are required. The CIE standard illuminant (A) was used as a white reference, because it simulates a tungsten-halogen light source (Wyszecki and Stiles, 1982). This illuminant is suitable, because our images of beef steaks were acquired with illumination provided by tungsten-halogen light bulbs.

Statistical Evaluation of the System

Our evaluation of system performance followed procedures described by Wheeler et al. (2002). They assessed performance of three instrumented tenderness prediction systems on the basis of progressive certification of steak sample tenderness in 10%

certification increments. Any steaks generating shear-force values greater than 4.53 kg (10.0 lbs) were classified as “tough.” The rest were “tender.” In the description that follows, “observed values” refers to the reference WBS values. “Predicted values” refers to the shear force predicted by the computer vision system.

Samples were first sorted and ordered on the basis of predicted values. For 10% certification levels, 10% of the steaks having the lowest predicted values were classified into a “certified tender” group and the remaining into a “not certified tender” group. The mean observed Warner-Bratzler shear-force values were compared for the two groups using a ‘t’ test for independent samples ($\alpha = 0.05$). First, equality of variance for the two groups was tested. If the variances were equal, a pooled variance estimate was used in the ‘t’ test. If the variances were not equal, Satterthwaite approximation was used to estimate the variance. A significant difference in mean observed shear-force values between the two groups would infer that the computer vision system had successfully sorted the “tender” from the “tough” samples at that certification level. Any “tough” sample (observed shear-force value greater than 4.53 kg) in the “certified tender” group was designated as an error. This procedure was repeated for certification levels up to 90%, in 10% increments. A 100% certification level signified classifying all samples as “tender” (without sorting). Thus, error at 100% certification level indicates the total number of “tough” carcasses.

Results & Discussion

Wavelet Textural Features

Each color band of the image was separately decomposed by db4 wavelet for five scales. At each scale, there were three orientations: horizontal, vertical, and diagonal. This procedure produced 15 decomposed images for each color band. In total, there were 45 decomposed images for a color image with three color bands. Energy features were calculated for each decomposed image. In addition, rotationally-invariant details were calculated by adding details (pixel-by-pixel) of three orientations at each scale, as outlined by Kim et al. (1998). Five rotationally-invariant detail images were produced for each color band. From these rotationally-invariant images of three color bands, 15 energy features and 15 correlation features were calculated. In total, 75 features were extracted per color image in RGB color space. Images were also transformed to HSI and CIE LAB color spaces. Similar features were extracted in these color spaces, bringing the total number features to 225 per image.

Regression Model

The number of features extracted from an image is higher than the total number of images used in the study. Also, several textural features were correlated with each other. Thus, a feature reduction step is necessary. A forward-selection method (significance level for entry = 0.20) was used in a linear regression model to identify a select number of features to predict 14-d WBS values. This procedure serves the purposes of both feature reduction and tenderness prediction. Twelve features were selected by the model (Table 7.2).

Table 7.2. Wavelet textural feature selected by the regression model.

No.	Feature Type	Color Bands	Scale	Orientation
1	Correlation	L*, a*	5	All ¹
2	Energy	b*	4	Diagonal
3	Correlation	S, V	2	All
4	Correlation	R, G	5	All
5	Correlation	L*, a*	4	All
6	Correlation	R, G	4	All
7	Correlation	S, V	5	All
8	Energy	b*	5	All
9	Energy	b*	4	All
10	Energy	L*	1	Horizontal
11	Energy	B	1	All
12	Energy	a*	3	Diagonal

¹All – rotationally-invariant details, obtained by summing details at all three orientations.

The first feature selected was the correlation between the wavelet-decomposed images at the 5th scale from the L* and a* bands. Correlation coefficient between the observed and predicted shear force values was 0.54. In comparison, 7-d WBS values predicted 14-d WBS values with a correlation coefficient of 0.53. Thus, predictions by textural features extracted from fresh images using the computer vision system (nondestructive) performed as well as 7-d WBS values (sample-destructive).

The beef industry is more interested in identifying “tough” carcasses than in predicting shear-force values. In our study, the limited number of tough samples (10) in the tested carcasses prevented evaluation of the system for classifying carcasses into

“tender” and “tough” categories. Therefore, the certification method described by Wheeler et al. (2002) was used to evaluate the system for sorting carcasses on the basis of nine tenderness certification levels.

Evaluation of the System

Certification results from the computer vision system are given in Table. 7.3. At all certification levels, there were significant differences between the observed WBS values for “certified tender” and “not certified tender” samples. This result indicates that the system is capable of sorting tough carcasses from tender carcasses with good accuracy.

Table 7.3. Statistical differences of the mean observed shear-force values between “certified tender” and “not certified tender” groups.

Certification Level	Observed Warner-Bratzler Shear Force, kg.		
	Certified Tender	Not Certified Tender	Difference
90	3.31	4.29	0.98*
80	3.29	3.91	0.62*
70	3.26	3.76	0.50*
60	3.20	3.72	0.52*
50	3.15	3.66	0.51*
40	3.11	3.60	0.49*
30	3.09	3.55	0.46*
20	3.00	3.51	0.51*
10	2.80	3.48	0.68*

*Mean observed shear-force values for “certified tender” and “not certified tender” were significantly different at $\alpha = 0.05$.

At the 100% certification level, no sorting was done. There were 10 tough samples in the data set (Fig. 7.6). Out of 10 tough samples, 8 were correctly identified in the “not certified tender” group at 90% certification level. Two tough carcasses were wrongly certified as tender. Note that the error in classification from lower certification levels is carried over to the error in higher certification levels. Thus, the error reported in Figure 7.6 is cumulative, as the certification levels increase from 10-90%. For example, one tough sample was incorrectly classified as “certified tender” at the 70% certification level. In addition to this error, another tough sample was incorrectly classified as “certified tender” at the 80% certification level, bringing the total number of misclassifications to 2. Below the 60% certification level, there were no errors.

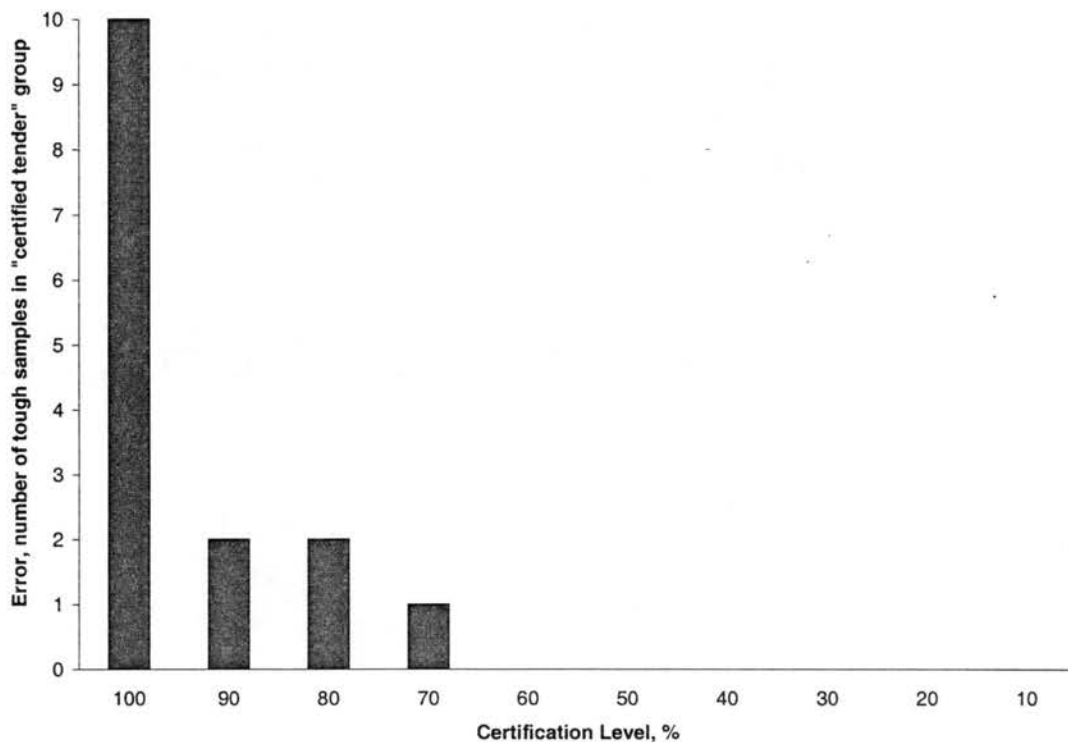


Figure 7.6. Error of the computer vision system at various certification levels.

Conclusion

Textural features were extracted in the wavelet domain using the 4th order Daubechies wavelet. A forward-selection method in the linear regression model was used for feature reduction and to predict 14-day aged, cooked-beef tenderness scores. Correlation between observed Warner-Bratzler shear-force values and predicted values was 0.54. In comparison, 7-d WBS values predicted 14-d WBS values with a correlation coefficient of 0.53. Thus, predictions by textural features extracted from fresh images using the computer vision system (nondestructive method) were as good as 7-d WBS values (sample-destructive method). At all tenderness certification levels, there were significant differences between the shear-force values for “certified tender” and “not certified tender” samples. These results show the ability of the system to accurately sort the steaks on the basis of cooked-beef tenderness. The system demonstrated the use of textural features extracted from images of fresh steaks to accurately predict aged, cooked-beef tenderness.

Acknowledgements

The authors thank the Food and Agriculture Products Research Initiative Program FRIP 2001, Oklahoma State University for supporting this study. The authors also thank Aaron Elam, Chad Carr, and Keith Charmasson, Animal Science Graduate Research Assistants for generating the shear-force data.

References

- AMSA. 1995. *Research guidelines for cookery, sensory evaluation, and instrumental tenderness measurements of fresh meat*. Savoy, IL.: National Live Stock & Meat Board, American Meat Science Association.
- Belk, K. E., J. D. Tatum, G. C. Smith, M. Goldberg, A. M. Wyle, and R. C. Cannell. 2001. Meat imaging system for palatability yield prediction. U.S. Patent No. 5793893.
- Benn, A., D. Barrett-Lennard, and P. J. Hay. 1998. Image analysis for meat. U.S. Patent No. 5793879.
- Boleman, S. L., S. J. Boleman, W. W. Morgan, D. S. Hale, D. B. Griffin, J. W. Savell, R. P. Ames, M. T. Smith, J. D. Tatum, T. G. Field, G. C. Smith, B. A. Gardner, J. B. Morgan, S. L. Northcutt, H. G. Dolezal, D. R. Gill, and F. K. Ray. 1998. National Beef Quality Audit-1995: survey of producer-related defects and carcass quality and quantity attributes. *Journal of Animal Science* 76(1): 96-103.
- Connolly, C., and T. Fleiss. 1997. A study of efficiency and accuracy in the transformation from RGB to CIELAB color space. *Image Processing, IEEE Transactions on* 6(7): 1046-1048.
- Gonzalez, R. C., and R. E. Woods. 1992. *Digital Image Processing*. Reading, M.A.: Addison-Wesley Publishing Company.
- Huang, Y., R. E. Lacey, L. L. Moore, R. K. Miller, A. D. Whittaker, and J. Ophir. 1997. Wavelet textural features from ultrasonic elastograms for meat quality prediction. *Transactions of the ASAE* 40(6): 1741-1748.

- IEEE Standard 610-4. 1990. *IEEE Standard Glossary of Image Processing and Pattern Recognition Terminology*. New York, NY.: IEEE Press.
- Jain, A. K. 1989. *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ.: Prentice Hall.
- Kim, N., D. Wilson, G. Rouse, and S. Udpa. 1998. Ultrasonic image texture analysis for characterizing intramuscular fat content of live beef cattle. *Ultrasonic Imaging* 20: 191-205.
- Li, J., J. Tan, and P. Shatadal. 1999. Discrimination of beef images by textural features. ASAE Paper No. 99-3158. St. Joseph, Mich.: ASAE.
- Li, J., J. Tan, and P. Shatadal. 2001. Classification of tough and tender beef by image texture analysis. *Meat Science* 57(4): 341-346.
- Livens, S., P. Scheunders, G. Van De Wouwer, and D. Van Dyck. 1997. Wavelets for texture analysis, an overview. In *Sixth International Conference on Image Processing and Its Applications*, 2: 581-585.
- Mckenna, D. R., D. L. Roebert, P. K. Bates, T. B. Schmidt, D. S. Hale, D. B. Griffin, J. W. Savell, J. C. Brooks, J. B. Morgan, T. H. Montgomery, K. E. Belk, and G. C. Smith. 2002. National Beef Quality Audit-2000: survey of targeted cattle and carcass characteristics related to quality, quantity, and value of fed steers and heifers. *Journal of Animal Science* 80(5): 1212-1222.
- Misiti, M., Y. Misiti, G. Oppenheim, and J. Poggi. 1996. Wavelet Toolbox for use with Matlab - User's Guide. Ver. 1. Natick, MA.: The Mathworks, Inc.

- Morgan, J. B., J. W. Savell, D. S. Hale, R. K. Miller, D. B. Griffin, H. R. Cross, and S. D. Shackelford. 1991. National beef tenderness survey. *Journal of Animal Science* 69(8): 3274-3283.
- Oppenheim, A. V., and R. W. Schaffer. 1999. *Discrete-Time Signal Processing*. 2nd ed. New Jersey, NJ.: Prentice Hall.
- Pollikar, R. 2002. The Wavelet Tutorial. Glassboro, NJ.: Department of Electrical and Computer Engineering, Rowan University. Available at:
<http://users.rowan.edu/~polikar/WAVELETS/WTtutorial.html>. Accessed on 21 July 2002.
- Shackelford, S. D., T. L. Wheeler, and M. Koohmaraie. 1997. Tenderness classification of beef: I. Evaluation of beef longissimus shear force at 1 or 2 days postmortem as a predictor of aged beef tenderness. *Journal of Animal Science* 75(9): 2417-2422.
- Strang, G., and T. Nguyen. 1996. *Wavelets and Filter Banks*. Wellesley, MA: Wellesley-Cambridge Press.
- Tong, A. K., D. J. Robinson, and T. Liu. 1998. Method and apparatus for using image analysis to determine meat and carcass characteristics. Canadian Patent No. 02263763.
- USDA. 1997. United States standards for grades of carcass beef. Effective date January 31, 1997. Washington, DC.: Livestock and Seed Division of the Agricultural Marketing Service. Available at:
<http://www.ams.usda.gov/lsg/stand/standards/beef-car.pdf>. Accessed on 9 February 2004.

- USDA. 2001. Procedures for approval and use of instrument augmentation systems for beef carcass ribeye measurement. Washington, DC.: Livestock and Seed Program, Agricultural Marketing Service.
- Van De Wouwer, G., P. Scheunders, and D. Van Dyck. 1999. Statistical texture characterization from discrete wavelet representations. *IEEE Transactions on Image Processing* 8(4): 592-598.
- Wheeler, T. L., D. Vote, J. M. Leheska, S. D. Shackelford, K. E. Belk, D. M. Wulf, B. L. Gwartney, and M. Koohmaraie. 2002. The efficacy of three objective systems for identifying beef cuts that can be guaranteed tender. *Journal of Animal Science* 80(12): 3315-3327.
- Wulf, D. M., S. F. O'Connor, J. D. Tatum, and G. C. Smith. 1997. Using objective measures of muscle color to predict beef longissimus tenderness. *Journal of Animal Science* 75(3): 684-692.
- Wyszecki, G., and W. S. Stiles. 1982. *Color Science. Concepts and Methods, Quantitative Data and Formulae*. 2nd ed. New York, NY.: John Wiley & Sons.

CHAPTER VIII
PREDICTING BEEF TENDERNESS USING NEAR-INFRARED
SPECTROSCOPY

S. Jeyamkondan, G. Kranzler, B. Morgan, and S. Rust

Abstract

A near-infrared spectral reflectance system was developed and tested online to predict 14-day aged, cooked-beef tenderness. A contact probe with a built-in tungsten-halogen light source supplied broadband light to the ribeye surface. Fiberoptics in the probe transmitted reflected light to a spectrometer with a spectral range of 400-2500 nm.

In the first phase, steak samples (n=292) were brought from packing plants to the lab and scanned with the spectrometer. After scanning, samples were vacuum-packaged and aged for 14 days. They were then cooked in an impingement oven to an internal temperature of 70° C. Slice shear-force values were recorded for tenderness reference.

In phase two, the spectrometer was modified for packing plant conditions. Spectral scans were obtained on-line on ribbed carcasses (n=276). A partial least squares regression model was developed to predict tenderness scores from spectral reflectance. In phase three, the developed model was validated by scanning carcasses (n=200) on-line. The predicted shear-force values and samples were sent to the U.S. Meat Animal

Research Center for third-party validation. At up to 70% tenderness certification levels, the system was able to successfully sort tough from tender carcasses.

Keywords: Beef tenderness, near-infrared spectroscopy, spectral reflectance, non-destructive evaluation.

Introduction

Beef tenderness is a primary consideration in consumer satisfaction. Despite its significance as a quality indicator, tenderness is not a current factor directly affecting product value for beef producers and packers. There is growing recognition that beef tenderness must be incorporated in the quality grading process if true value-based marketing is to take place. Developing instrumentation to predict tenderness was identified as a top priority for the beef industry by the National Cattlemen's Beef Association (NCBA, 2002). In the U.S. meat marketing system, beef products leave the packing plant at about three days postmortem. It takes approximately 14 days for the beef products to reach the consumer. The beef industry needs an instrument that can scan fresh meat at 2-3 days postmortem and predict ultimate 14-day cooked-beef tenderness. Major challenges stem from variability in the aging process and in cooking.

USDA Beef Quality Grading

The primary method used today to classify carcasses into various palatability groups is USDA quality grading (USDA, 1997). This method is employed because it is nondestructive and can be implemented online. The majority of U.S. beef falls into a narrow range of quality-grade variation. Roughly 80% of carcasses grade either USDA

Select or Low Choice (McKenna et al., 2002). Consumer dissatisfaction with the eating quality of beef generally relates to carcasses within this range. This match has led to questioning the economic value of the Quality grade. Graders tend to favor mid-range estimates instead of using the whole Quality grade range of U.S. Prime to U.S. Standard. Error rates for assignment of USDA Quality and Yield grades have been reported to be as high as 12%. In other words, one of eight carcasses is stamped with the wrong USDA Quality grade (J.W. Wise, 1997, USDA; personal communication).

Instrument Grading

Instrument grading has the potential to improve prediction of palatability by employing technology to sort carcasses into Quality groups. Three technologies that have been developed to determine beef tenderness are the U.S. Meat Animal Research Center (MARC) Slice Shear Force, the Colorado State University (CSU) BeefCam®, and the Colorimeter developed at South Dakota State University. A brief overview of these methods follows.

The MARC Slice Shear Force system uses belt-driven, automated equipment to facilitate sample processing (Shackelford et al., 1999). Shear-force values are obtained at three days postmortem and regressed to predict shear force at 14 days postmortem. This method is accurate at identifying beef that can be guaranteed tender. Because it is sample-destructive and relatively time-consuming, the system is not yet acceptable for on-line application. Perhaps the greatest drawback associated with this system is that it measures beef tenderness at early postmortem. It does not take into account variation in tenderness associated with postmortem aging.

Colorado State University (CSU), in collaboration with Hunter Associates Laboratory, Inc. (Reston, VA) developed the BeefCam[®] system to predict beef tenderness. The CSU BeefCam[®] is a video image analysis system that processes images of the ribeye to predict tenderness (Belk et al., 2001). This method meets the criteria of being nondestructive and relatively fast.

Wulf and Page (2000) developed a quality grade augmentation system that supplements current USDA Quality Grade Standards with Colorimeter readings ($L^*a^*b^*$), pH measurements, and hump height to predict beef tenderness.

In a study supported by the NCBA (Wheeler et al., 2002), these tenderness prediction technologies were tested. The following conclusions were reached:

- Indirect, non-invasive methods to predict meat tenderness that are based primarily on lean color are not sufficiently accurate to warrant current use.
- BeefCam[®] did not perform well in this study.
- The Colorimeter performed inconsistently, particularly on USDA Select carcasses.
- Direct measure of tenderness provided by Slice Shear Force resulted in a more accurate identification of “tender” beef carcasses. Because this method is sample-destructive and time-consuming, it appears that industry acceptance is unlikely.

A sufficiently accurate nondestructive instrument to predict aged beef tenderness continues to elude the beef industry.

Near-infrared Spectroscopy

Light reflected in the visible region of the spectrum gives an objective measurement of the color of food objects, whereas light reflected in the near-infrared (NIR) region contains information about food physical and chemical properties. Contradictory results have been reported in the literature regarding the use of NIR spectroscopy for predicting beef tenderness. Some studies have been conducted using laboratory spectrometers with special requirements for sample preparation. Others have used fiber-optic reflectance probes that can be more easily applied in a processing plant environment. In most studies, the NIR spectral scan and reference shear-force tenderness values were acquired at the same point in time. A few studies have used scans acquired at early postmortem to predict tenderness after aging. It is important to recognize these differences in order to understand the inconsistent results reported in the literature.

Pioneering Work

Use of NIR spectroscopy for predicting beef tenderness was pioneered by Mitsumoto et al. (1991). A limited number of samples was used in the study (11 steers, 6 muscles per steer). They collected spectral reflectance data (1100-2500 nm) on 3-d postmortem samples to predict 3-d Warner-Bratzler shear-force (WBS) values. Samples were prepared and placed in sample holders for scanning. A multiple linear regression equation with four selected wavelengths predicted 3-d WBS values with R^2 values of 0.68. This study showed promise and aroused broad interest in using NIR for predicting beef tenderness.

Matforsk Norwegian Research Group

A series of studies were conducted at the Matforsk Norwegian Food Research Institute. Hildrum et al. (1994) conducted a feasibility study with 10 carcasses (3 cows, 7 bulls) at three aging periods (1, 8, and 14 days). They acquired scans in the 1100-2500 nm range under controlled conditions requiring special sample preparation. Thirty spectral scans were used to build a two-factor principal component regression model with multiplicative scatter correction (MSC) data preprocessing. The model predicted WBS values with an R^2 value of 0.62. Note that 1-d spectra predicted 1-d WBS values, and 14-d spectra predicted 14-d WBS values.

In continuation of the work, Hildrum et al. (1995) used 40 carcasses (both cows and bulls) at three aging periods (1, 7, and 14 days). Thirty carcasses were chilled slowly. Ten carcasses were subjected to electric stimulation and rapid chilling. Spectra were collected from fresh samples and frozen/thawed samples. It is interesting to note that the NIR spectra collected from frozen/thawed samples predicted WBS values more accurately ($R^2 = 0.50$) than those collected fresh ($R^2 = 0.34$). The authors speculated that the freezing step increased the difference in structure between tender and tough samples. Note that the model predicted current WBS values rather than forecasting ultimate WBS values.

Rødboten et al. (2000) studied the feasibility of using early (4- and 26-hour) postmortem NIR spectroscopy to predict 2-d and 7-d WBS values from 127 carcasses. Spectra recorded at 4-h and 26-h postmortem exhibited some differences in the 1100-1900 nm range. The rigor mortis process starts at about 4-h postmortem and ends before 26 hours. The authors hypothesized that the difference between the two spectra collected

at the beginning and end of rigor mortis is related to the aging potential of the carcasses. However, they found that spectra differences did not contain essential information to predict final tenderness of aged samples. Partial least squares (PLS) models with MSC preprocessing of 4-h NIR spectra predicted 2-d and 7-d WBS readings with R^2 values of 0.25 and 0.27, respectively. Models based on 26-h spectra predicted 2-d and 7-d WBS readings with R^2 values of 0.36 and 0.37, respectively. They concluded that their study did not support the hypothesis that early postmortem NIR spectroscopy can be used as a predictor of final tenderness.

Rødbotten et al. (2001) used a spectrometer (950-1700 nm) equipped with a diode array detector to predict tenderness of 12 carcasses. Traditional NIR spectrometers use grating and require scanning to acquire complete spectra. Diode array-based spectrometers acquire spectra without scanning, making acquisition more rapid. The authors were able to acquire 150 spectra to cover the entire surface of a sample on both sides in 7 s. Four samples were collected from each carcass. NIR spectra and WBS readings were recorded at 2, 9, and 21 days postmortem. Separate models for different aging periods predicted current tenderness with R^2 value of 0.36-0.69. With regard to forecasting tenderness, a 2-d spectra model predicted 9-d WBS values with moderate accuracy ($R^2 = 0.52$). Less accuracy was reported for forecasting 21-d WBS values ($R^2 = 0.27$). Special sample preparation was not required for this spectrometer.

Teagasc Irish Research Group

Byrne et al. (1998) used a scanning spectrometer (750-1098 nm) equipped with a fiber optic probe to predict beef tenderness of 70 heifers. Due to lamp replacement during the course of this study, two separate models were developed using 20 and 50

samples, respectively. Results of the second set are described here. A 10-factor model based on 1-d spectra predicted 14-d WBS values with $R^2 = 0.68$. Models based on 2-, 7-, and 14-d spectra predicted less accurately ($R^2 = 0.20 - 0.45$). This result is intriguing, because 1-d spectra were able to predict 14-d WBS values more accurately than 14-d spectra. This result contradicts expectations, because a 1-d spectrum is forecasting tenderness that involves the variability of aging, whereas a 14-d spectrum predicts current tenderness. Even the prediction accuracies of models based on 1-d and 2-d spectra were considerably different. NIR readings and WBS values were collected on different steaks from the same carcasses. The authors suggested that the collecting NIR readings and WBS values from the same meat sample would improve the accuracy.

Continuing the above work, Venel et al. (2001) acquired NIR spectra (750-1098 nm) on *longissimus dorsi* (l.d.) and *semimembranosus* muscles from 75 animals. Following the recommendation made by Byrne et al. (1998), the site of spectral measurement was marked with a scalpel, and WBS values were measured at the same site. Spectra and WBS values were collected at 14-d post mortem. NIR spectroscopy was unable to predict tenderness for *semimembranosus* muscle. For l.d. muscle, R^2 values for tenderness prediction ranged from 0.27-0.33. When separate models were developed for each segregation of samples based on sex, grade, and pH, R^2 values of prediction ranged from 0.15-0.54.

Other Work

Park et al. (1998) collected NIR spectra (1100-2498 nm) and WBS values on frozen and thawed samples at either 7-d or 14-d post mortem on 119 samples. Special sample preparation was required. They developed a 6-factor PLS model to predict

tenderness with R^2 values of 0.67 and 0.63 for calibration (n=80) and validation (n=39) sets, respectively. Note that 7-d spectra predicted 7-d WBS values and 14-d spectra predicted 14-d WBS values. Hildrum et al. (1995) reported that the spectra collected on the samples that were frozen/thawed predicted WBS values more accurately than those collected on fresh samples. However, in industry application, samples will not be frozen and thawed.

Liu et al. (2003) collected NIR spectra (400-2498 nm) and WBS values from 24 carcasses at 2, 4, 8, 14, and 21 days postmortem. An 8-factor PLS model predicted WBS values with $R^2 = 0.49$. Note that the model did not forecast aged tenderness. When the data were segregated into different aging periods, prediction accuracy increased for 4-d and 21-d aging periods, while accuracy decreased for other aging periods.

Leroy et al. (2003) used a Fourier Transform NIR spectrometer to record spectra (833-2500 nm) from 189 samples using transmission and reflectance modes at 2-d and 8-d postmortem. Using the reflectance mode, the 2-d spectra model predicted 2-d and 8-d WBS readings with R^2 values of 0.25 and 0.19, respectively. With the transmission mode, a 2-d spectra model predicted 2-d and 8-d WBS readings with R^2 values of 0.41 and 0.15, respectively.

To summarize, several studies reported moderate to promising results in predicting current-status tenderness (Mitsumoto et al., 1991; Hildrum et al., 1994, 1995; Rødbotten et al., 2001; Park et al., 1998; Liu et al., 2003). Only one study (Byrne et al., 1998) reported success in forecasting tenderness. However, another study from the same research group (Venel et al., 2001) reported failure in predicting current tenderness.

Three studies reported failure in forecasting tenderness (Rødboten et al., 2000, 2001; Leroy et al., 2003).

Objectives

The overall objective of this study is to investigate the feasibility of using NIR spectral reflectance to predict tenderness in a packing-plant environment. Specific objectives were:

- Develop a portable system to collect near-infrared spectral reflectance values from fresh meat at 3-d postmortem in a packing-plant environment.
- Develop chemometric models to relate spectral reflectance to 14-d shear-force tenderness.
- Evaluate the system in a packing-plant environment.

Materials & Methods

Meat Samples

Beef ribeye roll samples (n=768) were collected from two regional packing plants. Quality grades of the tested carcasses were 50% USDA Select and 50% Low Choice. This project was conducted in three phases. In Phase I (laboratory scanning), carcasses (n=292) were selected at the USDA grading station (approximately 48 h postmortem). Carcasses identified and selected were individually routed off onto separate rails by USDA quality grade. Approximately 100 carcasses per grade per plant were selected. Carcass grade data factors such as ribeye area, lean maturity, skeletal maturity, marbling score, and quality grade as evaluated and stamped by USDA grader

were collected. Hot carcass weight and carcass identification numbers were recorded from plant tags.

Following grade data collection, all carcasses were fabricated. Individually identified ribeye rolls (IMPS# 112) were collected, vacuum-packaged, and packed into ice chests with ice packs for transportation from the packing plant to the Oklahoma State University (OSU) Food and Agricultural Products Research Center. At approximately 72 h postmortem, a one-inch (2.54 cm) steak was cut from the anterior end of each ribeye roll, individually identified, and allowed to “bloom” for 30 minutes. Sample steaks were moved to the OSU Biosystems & Agricultural Engineering Machine Vision Laboratory for spectral scanning under controlled conditions. Following scanning, steaks were vacuum-packaged, aged for 14 days (1.0°C), and then frozen (-2.0°C) for later shear-force measurements. After thawing, steaks were cooked, chilled for 24 hours, and then slice shear-force values were measured. Slice shear-force measurement (described later) was used as tenderness reference.

In Phase II, carcasses (n=276) from two commercial plants were selected at the grading stand and routed onto separate rails by USDA quality grade. In-plant scanning was performed on each carcass (n=100, USDA Choice; n=100, USDA Select). Carcass grade and shear data were collected, and Phase I procedures were repeated. The concluding portion of the project (Phase III) consisted of third-party verification that included in-plant scanning and delivering the scanned ribeye steaks (n=200), along with predicted tenderness categories, to the U.S. Meat Animal Research Center in Clay Center, Nebraska for slice shear-force measurements. These tenderness ratings were established

from the in-plant spectral scans obtained from the two cooperating beef processing facilities.

NIR Spectrometer

The spectrometer used in this study (FieldSpec Pro JR, Analytical Spectral Devices, Inc.) scans in the visible and NIR regions (400-2500 nm). For in-plant operation, the spectrometer is secured in a backpack, with a notebook computer supported in view of the operator (Fig. 8.1a). The fiberoptic cable from the spectrometer terminates in a contact probe that projects broadband lighting and positions the cable to collect the light reflected from the beef surface (Fig. 8.1b).

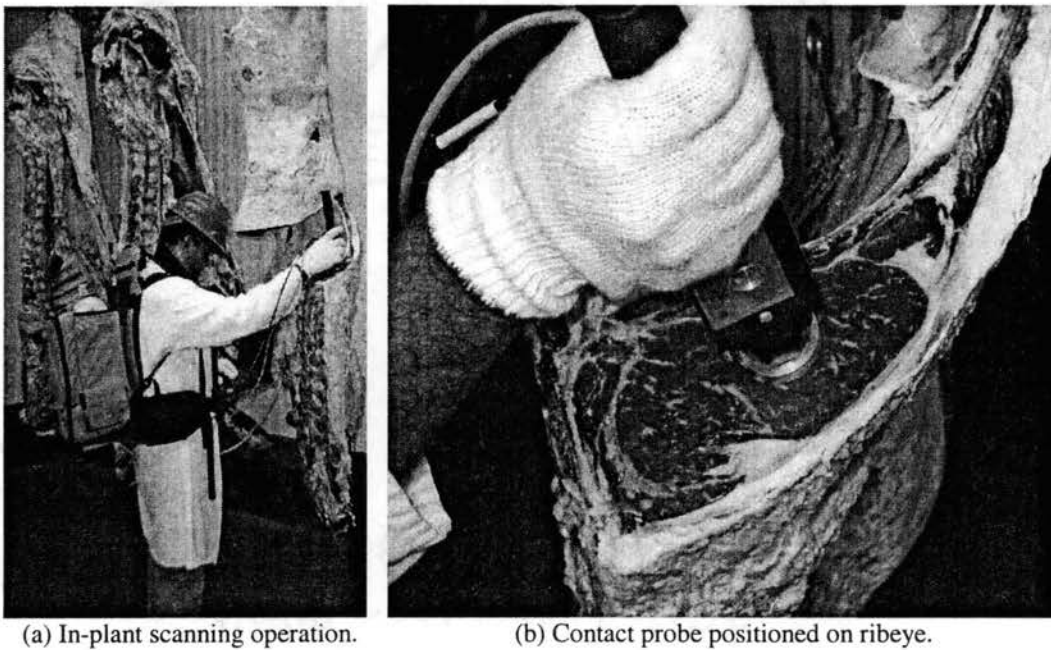


Figure 8.1. Scanning a sample in a packing plant using a portable spectrometer with contact probe.

Reflectance Measurement

Optimization of integration time was accomplished by periodic placement of the contact probe on a white reference plate (Spectralon[®] Diffuse Reflectance Targets,

LabSphere, Inc.). For a given scan, 10 spectra were collected consecutively and averaged to minimize electronic noise. Three spectra were collected at three locations near the lateral end of the ribeye, in an effort to avoid connective tissue. The median of the three spectra was calculated and saved as the reflectance spectrum for that beef sample. Calculating the median avoided the effect of outliers related to a thick marbling spot or connective tissue.

Shear-Force Measurement

Traditionally, Warner-Bratzler shear-force values have been used as the tenderness reference. This procedure involves coring steaks parallel to muscle fibers to collect cylindrical samples for shearing. Coring angle and location of core samples were not standardized, leading to limited repeatability. Precision of reference values is critical in developing reliable models. A new tenderness reference method known as “slice shear” has been developed by the U.S. Meat Animal Research Center (Shackelford et al., 1999). In their method, a larger sample is sliced at a fixed location on the ribeye, at a 45° angle. This procedure produces better repeatability than the WBS method. The slice shear method was used as the tenderness reference in this study.

Steaks for slice shear-force (SSF) assessment were thawed for 24 hours at 1 to 2°C and cooked on a belt-fed impingement oven (Model 1132-000-A, Lincoln Foodservice Products, Inc.). Preliminary test cooking was done to determine appropriate cooking times to reach a 71°C internal temperature. Slice shear force was measured after the cooked steaks were allowed to chill for 24 hours at 4°C. Using the procedures outlined by Shackelford et al. (1999), a first cut was made 1 cm from the lateral end of the cooked steak. The second cut was made at 5 cm from the first cut. The slice shear

sample was removed at an angle of 45° using a knife with two parallel blades separated by a 1-cm space. This procedure generated a cooked meat sample measuring 5 cm in length by 1 cm in thickness and 2.5 cm in width. This sample location was selected so that limited connective tissue would be located within the slice shear sample. Slice shear force was measured using a flat, blunt-end blade (slice shear attachment) attached to an Instron Universal Testing Machine (Instron Corp, Canton, MA). Force required to shear the muscle fibers of the slice was recorded as the reference slice-shear force. Higher slice shear-force values indicated “tougher” beef.

Model Development

This is a multivariate data analysis problem. Each spectrum has 2,150 data points or independent variables. Slice shear force is the dependent variable. A dimensionality reduction technique must be employed to avoid over-fitting. We used a partial least square (PLS) regression. PLS produces new features that are linear combinations of original spectral data points such that the new factors are not correlated and explain most of the variation in both dependent and independent variables (CAMO, 1998).

Absorbance spectra in the region of 400-1650 nm were used to predict slice shear force. Signal strength from reflections beyond 1650 nm was below threshold. The model was developed using Unscrambler software (Camo, Inc.). Cross-validation procedures (Esbension, 2001) were employed to select the number of PLS factors to include in the models.

Evaluation of Statistical Model

Our evaluation of system performance followed procedures described by Wheeler et al. (2002). They assessed performance of three instrumented tenderness prediction systems on the basis of progressive certification of steak sample “tenderness” in 10% certification increments. We classified steaks with 14-d slice shear force greater than 25 kg as “tough” and the rest as “tender.” In the description that follows, “observed values” refers to the reference slice shear-force values. “Predicted values” refers to the 14-d shear force predicted by the spectral reflectance system.

Samples were first sorted and ordered on the basis of predicted values. For 10% certification levels, 10% of the steaks having the lowest predicted values were classified into a “certified tender” group and the remaining into a “not certified tender” group. The mean observed slice shear-force values were compared for the two groups using a ‘t’ test for independent samples ($\alpha = 0.05$). First, equality of variance for the two groups was tested. If the variances were equal, a pooled variance estimate was used in the one-tailed ‘t’ test. One-tailed test was conducted, because the mean shear-force value of “not certified tender” group was expected to be greater than that of “certified tender” group. If the variances were not equal, Satterthwaite approximation was used to estimate the variance. A significant difference in mean observed shear-force values between the two groups would indicate that the spectral reflectance system had successfully sorted the tender from the tough samples at that certification level. Any tough sample (observed 14-d SSF ≥ 25 kg) in the “certified tender” group was an error. This procedure was repeated for certification levels up to 90%, in 10% increments. A 100% certification level signified classifying all samples as tender (without sorting).

Results & Discussion

Meat Samples

The system was developed and tested using carcasses with a wide range of hot carcass weights and USDA yield grades (Table 8.1). Most of the carcasses were of “A” maturity and graded either “Select” or “Choice.” This sample selection is in agreement with typical packing-pant grade distribution as reported in the recent National Beef Quality Audit – 2000 (McKenna et al., 2002).

Table 8.1. Carcass and shear data summary of all samples (n=768).

Variable	Mean	SD	Minimum	Maximum
Hot Carcass Weight, lb	760.8	250.1	520.4	1001.7
USDA Yield Grade	2.5	0.8	0.6	4.6
Selected Maturity ^a	162	12	135	240
Lean Maturity ^a	167	28	130	270
Marbling Score ^b	492	79	230	860
USDA Quality Grade ^c	670	64	510	830
Slice Shear Force, kg	18.15	4.14	9.87	39.87

^aMaturity: 100=A⁰⁰, 200=B⁰⁰, etc.

^bMarbling: 200=Practically Devoid⁰⁰, 300=Traces⁰⁰, 400=Slight⁰⁰, 500=Small⁰⁰, 600=Modest⁰⁰, etc.

^cUSDA Quality Grade: 500=Standard⁰⁰, 600=Select⁰⁰, 700=Choice⁰⁰, 800=Prime⁰⁰

NIR Readings

In general, tough meat absorbed more light than tender meat. This result is in agreement with previous studies (Rødbotten et al., 2000; Leroy et al., 2003). Beyond 1,400 nm, the signal level is low due to water absorption. The reflectance (R) spectrum is converted to an absorbance spectrum by the log (1/R) transformation. This

transformation is commonly employed to linearize the relationship between the concentration of an absorbing compound and the absorption spectrum (Hruschka, 2001). A typical absorbance spectrum is shown in Fig. 8.2.

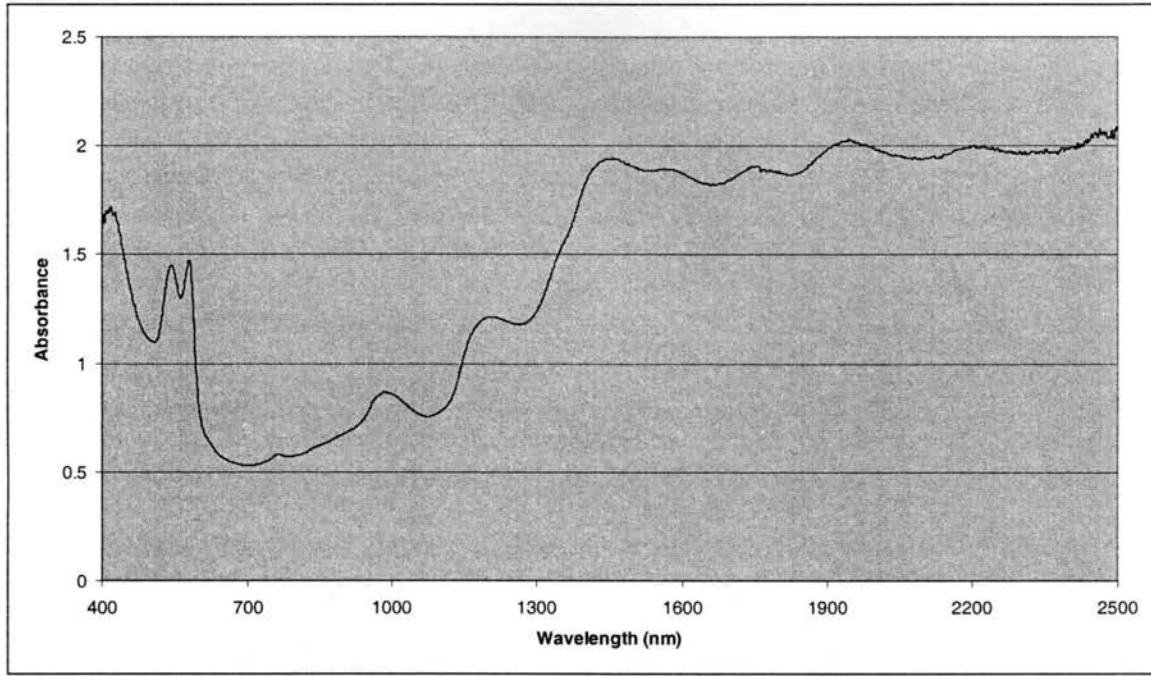


Figure 8.2. Typical absorbance spectrum from beef ribeye.

The spectra recorded in the laboratory were slightly different from those recorded in the packing plant due to different environmental conditions. Therefore, the model developed from Phase II samples acquired in the packing plant was used to predict slice shear-force values for samples in Phase III. Phase II and III samples were labeled as calibration and validation samples, respectively. A 1-factor PLS model was found to be optimum. The correlation coefficient between the observed and predicted values was 0.34. Standard error of calibration and prediction were both 4.17 kg. Note that slice shear-force values greater than 25 kg are considered “tough” carcasses.

A low correlation coefficient between the observed and predicted slice shear-force values indicated that the system did not predict specific tenderness values with high accuracy. However, the beef industry is more interested in identifying tough carcasses than in predicting exact shear-force values. A categorical model could be developed to classify carcasses into tender ($SSF < 25$ kg) or tough ($SSF \geq 25$ kg). By this criterion, only 6.8% of the carcasses tested in this study were tough. It follows, that an instrument designed to automatically classify all carcasses as “tender” would be 93.2% accurate in this study. This anomaly stems from the low number of tough carcasses in the sample population. Wheeler et al. (2002) faced this same problem when they evaluated three instruments for predicting beef tenderness. They developed a “tenderness certification” method to evaluate and compare the instruments. Instead of using one cutoff point ($SSF = 25$ kg) to classify a carcass as tender or tough, they used nine cutoff points. The samples were sorted, based on instrument prediction. At 10% certification level, 10% of samples with the lowest predicted shear-force values were certified as tender and the remaining 90% as tough. If the instrument sorted the carcasses successfully, there would be a significant difference in the statistical means of the actual (observed, not predicted) shear-force values. This procedure was continued up to 90% certification level, in 10% increments.

Table 8.2 shows the mean shear-force values for “certified tender” and “not certified tender” groups at various certification levels. Significant differences at all certification levels would indicate that the system successfully sorted the carcasses into two certified tenderness groups. Table 8.3 shows the certification table for Phase III (validation) samples. Significant differences between the mean shear-force values

between the two groups were observed at or below 70% certification levels. Performance of the system in the validation data set is slightly lower than that in the calibration data set.

Table 8.2. Mean shear-force values of Certified Tender and Not Certified Tender groups for Phase II (calibration) samples.

Percentage Certification Level	Slice Shear Force (kg)		
	Certified Tender	Not Certified Tender	Difference
90	18.21	20.88	2.67*
80	17.89	20.87	2.98*
70	17.65	20.44	2.79*
60	17.50	19.96	2.46*
50	17.26	19.71	2.45*
40	17.21	19.32	2.11*
30	16.70	19.24	2.54*
20	16.35	19.01	2.66*
10	16.03	18.75	2.72*

*Difference between “certified tender” and “not certified tender” significant ($P < 0.05$).

One factor affecting comparative performance may be the different cooking modes used for calibration and validation steak samples. As mentioned earlier, SSF values were generated for calibration samples in-house at Oklahoma State University. Calibration samples were cooked in a belt-fed traditional impingement oven. Cooking time was about 10 minutes. SSF values for validation samples were generated by MARC. The MARC personnel used a belt grill operating at a high temperature for the “very rapid” cooking protocol described in Shackelford et al. (1999). Cooking time was 4.4 minutes. Shackelford et al. (1999) observed that “very rapid” cooking resulted in

more surface crust formation and drip loss. Differences in cooking could have affected comparative slice shear-force value, compromising analysis of results.

Table 8.3. Mean shear-force values of Certified Tender and Not Certified Tender groups for Phase III (validation) samples.

Percentage Certification Level	Slice Shear Force (kg)		
	Certified Tender	Not Certified Tender	Difference
90	14.60	14.80	0.20
80	14.62	14.64	0.03
70	13.99	16.09	2.09*
60	13.71	15.99	2.28*
50	13.44	15.80	2.36*
40	13.10	15.63	2.53*
30	13.01	15.31	2.30*
20	13.39	14.93	1.54*
10	13.05	14.80	1.75*

*Difference between “certified tender” and “not certified tender” significant ($P < 0.05$).

Nevertheless, significant differences were observed at or below 70% certification levels. This result indicates that the system successfully sorted the tough from tender carcasses at or below 70% certification levels. Practical implications to the beef industry are that, if the top 30% of carcasses sorted “tough” by this system are removed, the remaining 70% can be sold as “guaranteed tender” for premium markets such as restaurants.

Conclusions

An NIR spectral reflectance system with contact probe was developed and evaluated on-line. The contact probe provided stable broadband light and fixed the geometry of light and fiberoptic probe in relation to the meat surface. Spectral reflectance values were collected at 3-d postmortem and were used to forecast 14-d slice shear-force tenderness values. Low correlation between the observed and predicted slice shear-force values indicated that the system did not predict exact tenderness values with high accuracy. However, the system was able to sort carcasses into tenderness categories with high accuracy. At up to 70% certification levels, the system successfully sorted carcasses into two certified tenderness categories. Implications to the beef industry are that 70% of the “certified tender” carcasses could be sold as “guaranteed tender” to premium markets.

Acknowledgements

This study was funded by the National Cattlemen’s Beef Association, the Oklahoma Beef Industry Council, and Oklahoma State University Food and Agriculture Products Research & Technology Research Initiative Program 2002. The authors would also like to thank Dr. T.L. Wheeler and personnel at the U.S. Meat Animal Research Center for generating slice shear-force values for study validation.

References

- Belk, K. E., J. D. Tatum, G. C. Smith, M. Goldberg, A. M. Wyle, and R. C. Cannell. 2001. Meat imaging system for palatability yield prediction. U.S. Patent No. 5793893.
- Byrne, C. E., G. Downey, D. J. Troy, and D. J. Buckley. 1998. Non-destructive prediction of selected quality attributes of beef by near-infrared reflectance spectroscopy between 750 and 1098 nm. *Meat Science* 49(4): 399-409.
- CAMO. 1998. The Unscrambler User Guide. Ver. Corvallis, OR.: Camo, Inc.
- Esbenson, K. H. 2001. *Multivariate data analysis - in practice. An introduction to multivariate data analysis and experimental design*. 5th ed. Corvallis, OR.: Camo, Inc.
- Hildrum, K. I., T. Isaksson, T. Naes, B. N. Nilsen, M. Rødbotten, and P. Lea. 1995. Near infrared reflectance spectroscopy in the prediction of sensory properties of beef. *Journal of Near Infrared Spectroscopy* 3: 81-87.
- Hildrum, K. I., B. N. Nilsen, M. Mielnik, and T. Naes. 1994. Prediction of sensory characteristics of beef by near-infrared spectroscopy. *Meat Science* 38(1): 67-80.
- Hruschka, W. R. 2001. Chapter 3: Data Analysis: Wavelength selection methods. In *Near-Infrared Technology in the Agricultural and Food Industries.*, P. Williams and K. Norris., eds. St. Paul, MN.: American Society of Cereal Chemist, Inc.
- Leroy, B., S. Lambotte, O. Dotreppe, H. Lecocq, L. Istasse, and A. Clinquart. 2003. Prediction of technological and organoleptic properties of beef Longissimus thoracis from near-infrared reflectance and transmission spectra. *Meat Science* 66(1): 45-54.

- Liu, Y., B. G. Lyon, W. R. Windham, C. E. Realini, T. D. D. Pringle, and S. Duckett. 2003. Prediction of color, texture, and sensory characteristics of beef steaks by visible and near infrared reflectance spectroscopy. A feasibility study*. *Meat Science* 65(3): 1107-1115.
- Mckenna, D. R., D. L. Roebert, P. K. Bates, T. B. Schmidt, D. S. Hale, D. B. Griffin, J. W. Savell, J. C. Brooks, J. B. Morgan, T. H. Montgomery, K. E. Belk, and G. C. Smith. 2002. National Beef Quality Audit-2000: survey of targeted cattle and carcass characteristics related to quality, quantity, and value of fed steers and heifers. *Journal of Animal Science* 80(5): 1212-1222.
- Mitsumoto, M., S. Maeda, T. Mitsuhashi, and S. Ozawa. 1991. Near-infrared spectroscopy determination of physical and chemical characteristics in beef cuts. *Journal of Food Science* 59(6): 1493-1496.
- NCBA. 2002. Beef Update. Meeting Summary. National Beef Instrument Assessment Plan II: Focus on Tenderness. Centennial, CO: National Cattlemen's Beef Association. Available at:
<http://www.beefboard.org/documents/beef%20update.pdf>. Accessed on 9 February 2004.
- Park, B., Y. R. Chen, W. R. Hruschka, S. D. Shackelford, and M. Koochmaraie. 1998. Near-infrared reflectance analysis for predicting beef longissimus tenderness. *Journal of Animal Science* 76(8): 2115-2120.
- Rødbotten, R., B.-H. Mevik, and K. I. Hildrum. 2001. Prediction and classification of tenderness in beef from non-invasive diode array detected NIR spectra. *Journal of Near Infrared Spectroscopy* 9: 199-210.

- Rødboten, R., B. N. Nilsen, and K. I. Hildrum. 2000. Prediction of beef quality attributes from early post mortem near infrared reflectance spectra. *Food Chemistry* 69(4): 427-436.
- Shackelford, S. D., T. L. Wheeler, and M. Koohmaraie. 1999. Tenderness classification of beef: II. Design and analysis of a system to measure beef longissimus shear force under commercial processing conditions. *Journal of Animal Science* 77(6): 1474-1481.
- USDA. 1997. United States standards for grades of carcass beef. Effective date January 31, 1997. Washington, DC.: Livestock and Seed Division of the Agricultural Marketing Service. Available at: <http://www.ams.usda.gov/lsg/stand/standards/beef-car.pdf>. Accessed on 9 February 2004.
- Venel, C., A. M. Mullen, G. Downey, and D. J. Troy. 2001. Prediction of tenderness and other quality attributes of beef by near infrared reflectance spectroscopy between 750 and 1100 nm; further studies. *Journal of Near Infrared Spectroscopy* 9: 185-198.
- Wheeler, T. L., D. Vote, J. M. Leheska, S. D. Shackelford, K. E. Belk, D. M. Wulf, B. L. Gwartney, and M. Koohmaraie. 2002. The efficacy of three objective systems for identifying beef cuts that can be guaranteed tender. *Journal of Animal Science* 80(12): 3315-3327.

Wulf, D. M., and J. K. Page. 2000. Using measurements of muscle color, pH, and electrical impedance to augment the current USDA beef quality grading standards and improve the accuracy and precision of sorting carcasses into palatability groups. *Journal of Animal Science* 78(10): 2595-2607.

CHAPTER IX
PREDICTING BEEF TENDERNESS USING
X-RAY ATTENUATION PROPERTIES

S. Jeyamkondan, G.A. Kranzler, P. Weckler, T. Bowser

Abstract

Tenderness is a critical factor in consumer perception of beef palatability. The objective of this study is to evaluate the feasibility of using X-ray attenuation properties of beef to predict aged, cooked-beef tenderness. Our digital soft X-ray imaging system consisted of a variable-energy X-ray source (10-50 kV; 0-1 mA), a photodiode array coupled with phosphor scintillation material, a 12-bit digitizer, and a computer. System calibration was executed and a regression model was developed to predict incident X-ray intensity based on voltage and current settings.

Beef steak samples (n = 378) were obtained from regional packing plants. X-ray images of steaks were acquired at voltages of 30, 35, 40, and 45 kV at 1 mA current and 460 ms integration time. After imaging, the steaks were aged for 14 days. Slice shear-force measurements were collected as the tenderness reference. R-value (ratio of low-energy mass attenuation coefficient to high-energy mass attenuation coefficient) was calculated pixel-by-pixel. Mean and standard deviation of R-value images were extracted

as attenuation features. These features were then related to 14-d shear-force tenderness scores.

No suitable model could be developed to predict tenderness. Enzymatic activity during aging is known to produce chemical changes that aid tenderization. Molecular composition does not change during this aging process. X-ray attenuation is affected by molecular composition of a material, but not by chemical structure. This study does not support the hypothesis that the X-ray attenuation property can be used to predict beef tenderness.

Keywords: X-ray imaging, mass attenuation coefficient, dual-energy X-ray absorption, R-value, beef tenderness.

Introduction

X-rays in the range of 10-50 keV (kilo electron-volts) are about 10,000 times more energetic than visible light. Due to this high energy content, the photons can penetrate most materials. Resistance to penetration or degree of attenuation varies with different materials. This attenuation property can be used to differentiate between materials on the basis of composition, thickness, and density. This relationship makes X-rays ideal for nondestructive quality evaluation of food products and detection of foreign objects.

Dual-energy X-ray absorption (DEXA) has been successfully used for body composition analysis in humans and animals. Theory and applications of DEXA have been demonstrated by Buzzell and Pintauro (2004). Mitchell et al. (1997) investigated DEXA for composition analysis of three-rib sections of beef using a commercial densitometer designed for medical imaging. DEXA measurements predicted the lipids,

protein, and fat content of ribs, with correlation coefficients ranging from 0.85 to 0.94. Continuing the work, Mitchell et al. (1998a) used DEXA to study composition of pork carcasses. Mitchell et al. (1998b, 1998c, 2001) used DEXA for body composition analysis and to measure bone mineral content and density in live pigs.

Brienne et al. (2001) used DEXA for assessment of meat fat content. They used a commercial densitometer that employs 24 sodium iodide detectors to obtain images at 44 and 70 keV at 0.2 mA. They positioned a series of 12 Plexiglass plates under the X-ray source tube to attenuate the X-ray photons and avoid detector saturation. Then, they estimated incident intensity, I_0 , by accounting for the Plexiglass absorption. DEXA predicted various chemical components with R^2 values ranging from 0.70 to 0.97.

A patent application has been filed in New Zealand covering evaluation of meat properties (including tenderness) using X-ray absorption. The patent disclosure indicates that the invention does not use an objective test of meat tenderness for comparison with X-ray measurements. Their “estimated tenderness” is based on unidentified visual features of the meat (Murray, 2001).

The objective of this study is to evaluate the feasibility of using X-ray attenuation properties to predict beef tenderness.

Materials & Methods

Principles of X-rays

When a high voltage is applied between the cathode and anode of an X-ray tube, electrons from the cathode accelerate toward the target anode. When the fast-moving stream of electrons is suddenly decelerated by the anode, X-rays are produced. X-rays

can be characterized by quantity and quality. Quantity refers to the number of the photons, whereas quality refers to the energy of the photons. Intensity is the product of the number of photons and associated energy. Thus, X-ray intensity depends on photon numbers and energy levels (Curry et al., 1990).

Quantity of X-rays can be increased by increasing X-ray tube current (mA). The X-ray tube current refers to the rate of electrons flowing between the two electrodes. Quality of X-rays can be increased by raising the peak voltage between the two electrodes. If a 50 kVp (peak kilovoltage) is applied between the two electrodes, the maximum energy achieved by the X-rays is 50 keV. In implementation, X-ray tubes actually produce a spectrum (range of wavelengths/energies) of X-rays.

In the atomic structure of a material, X-ray photons interact with the orbital electrons. X-rays with energy levels less than 50 keV interact with matter either by scattering or by absorption. Scattering is the undesired form of interaction. In scattering, X-ray photons are deflected by the orbital electrons and end up being detected in the “wrong” location on the detector. This creates a “fog” in an X-ray image (Curry et al., 1990).

Photoelectric absorption is the more desired form of interaction, from the standpoint of image quality. When an X-ray photon interacts with inner orbits such as K-, L-, or M-shells, the photon loses all its energy by ejecting the electron from the orbit. The excess energy is transmitted as the kinetic energy of the ejected electron. Because the photon loses all its energy, it disappears. The result of this interaction is termed photoelectric absorption (Curry et al., 1990).

Beer's law states that the intensity of X-rays is attenuated exponentially due to interaction with matter (Curry et al., 1990):

$$I = I_o e^{-\mu_m \rho t} \quad (9.1)$$

Where: I = transmitted X-ray intensity,

I_o = incident X-ray intensity,

μ_m = mass attenuation coefficient (cm^2/g),

ρ = material density (g/cm^3),

t = material thickness (cm).

It is important to note that the above relationship is valid only for monochromatic X-rays, i.e., the mass attenuation coefficient is defined only for X-ray photons having the same energy level. The mass attenuation coefficient is a material property and therefore does not vary with the thickness of the material. It is constant for X-rays of a given energy level.

X-ray Imaging System

A schematic of our X-ray imaging system is shown in Figure 9.1. The X-ray source tube (Model XTF-5011 Oxford Instruments X-ray Technology, Inc.) consists of a cathode and anode in an evacuated envelope. The tube includes a beryllium window (atomic number 4), which has a low attenuation coefficient. The beryllium window filters out very-low energy X-rays (less than 5 keV). The X-rays leaves the tube as an elliptic spot and spreads out. A cooling fan is placed above the tube.

The X-ray camera (Model Shad-o-BoxTM 1024 Rad-icon Imaging Corp.) consists of a two-dimensional photodiode array of size 1024 x 1024. The imaging area is 50 mm

x 50 mm. Center-to-center distance between pixels is $48 \mu\text{m}$. Above the photodetector array, a $\text{Gd}_2\text{O}_2\text{S}$ scintillator screen converts the X-rays to visible light that is detected by the photodetector array. A graphite window shrouds the scintillator to shield against ambient light. The sensors are optimized to detect X-rays in the range of 10 to 50 keV.

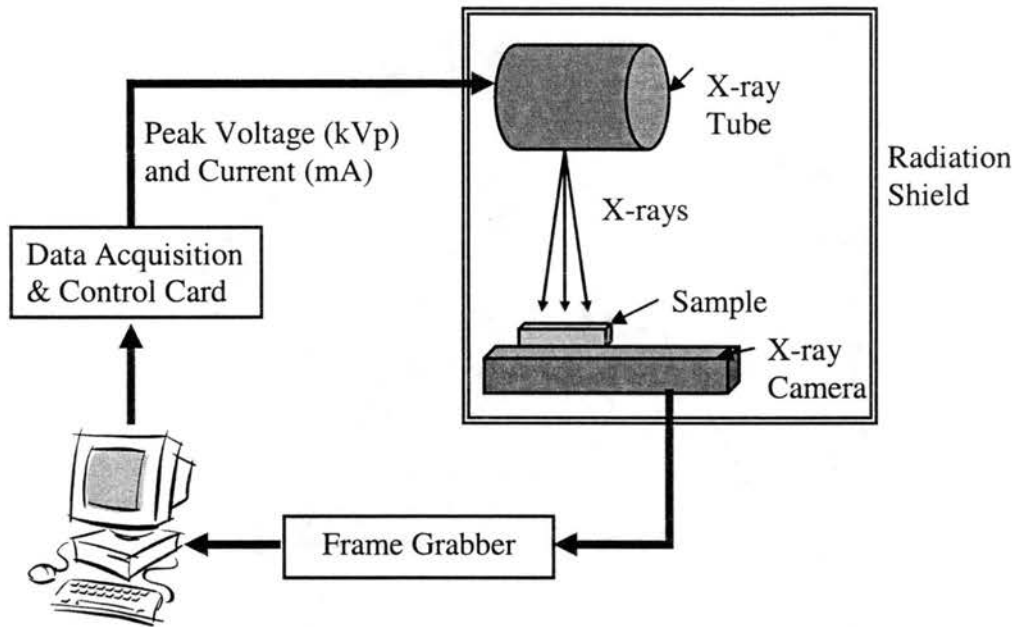


Figure 9.1. Schematic of X-ray imaging system.

The analog signal from the photodiode array is transmitted to a 12-bit frame grabber through a parallel interface. Maximum frame rate is just above 2 frames/s. The minimum integration time to acquire an image frame is 460 ms.

The camera/sensor is set on a height-adjustable holder, with the X-ray source tube fixed above. The sample object is placed between the tube and the camera. Input voltage and current supplied to the tube are controlled by a data acquisition and control card (DAQ-801, Omega Engineering Inc.). Voltage can be varied from 10-50 keV. Current is adjustable from 0.1 to 1 mA.

The source tube and camera are boxed in a 500 x 500 x 500 mm aluminum X-ray shield. Lead sheeting (6.35-mm) lines the interior. Two Plexiglass side windows allow for visual inspection. A safety switch in the door automatically cuts power to the X-ray tube when the door is opened.

Samples

Beef ribeye steaks (n=378) were collected from two regional packing plants. After selecting carcasses at the USDA grading station (approximately 48-h postmortem), carcass grade data factors were collected. Following grade data collection, all carcasses were fabricated. Individually identified ribeye rolls (IMPS# 112) were collected, vacuum-packaged, and packed into ice chests for transportation back to the Oklahoma State University (OSU) Food and Agricultural Products Research Center. At approximately 72-h postmortem, a one-inch steak (2.54 cm) was cut from the anterior end, individually identified, and allowed to “bloom” for 30 minutes. Trays were moved to the OSU Biosystems & Agricultural Engineering Machine Vision Laboratory for X-ray imaging.

X-ray Imaging

X-ray images of steaks were acquired at voltages of 30, 35, 40, and 45 kV at 1 mA current and 460 ms integration time. All four images of a steak were acquired without altering the position of the steak sample. After imaging, the steaks were vacuum-packaged, aged for 14 days, and then frozen (-2.0° C) for later shear-force tenderness measurements.

Slice Shear-Force Measurement

Traditionally, Warner-Bratzler shear force values have been used as the tenderness reference. This procedure involves coring steaks parallel to muscle fibers to collect cylindrical samples for shearing. Coring angle and location of core samples were not standardized, leading to limited repeatability. Precision of reference values is critical in developing reliable models. A new tenderness reference method known as “slice shear” has been developed by the U.S. Meat Animal Research Center (Shackelford et al., 1999). In their method, a larger sample is sliced at a fixed location of the ribeye, at a 45° angle. This procedure produces better repeatability than the WBS method. The slice shear method was used as the tenderness reference in this study.

Steaks for slice shear-force (SSF) assessment were thawed for 24 hours at 1 to 2° C and cooked on a belt-fed impingement oven (Model: 1132-000-A Lincoln Foodservice Products, Inc.). Preliminary test cooking was done to determine appropriate cooking times to reach a 71°C internal temperature. Slice shear force was measured after the cooked steaks were allowed to chill for 24 hours at 4°C. Using the procedures outlined by Shackelford et al. (1999), a first cut was made 1 cm from the lateral end of the cooked steak. The second cut was made at 5 cm from the first cut. The slice shear sample was removed at an angle of 45° using a knife with two parallel blades separated by a 1-cm space. This procedure generated a cooked meat sample measuring 5 cm in length by 1 cm in thickness and 2.5 cm in width. This sample location was selected so that limited connective tissue would be located within the slice shear sample. Slice shear force was measured using a flat, blunt-end blade (slice shear attachment) attached to an Instron Universal Testing Machine (Instron Corp, Canton, MA). Force required to shear

the muscle fibers of the slice was recorded as the reference slice shear force. Higher SSF values indicated “tougher” beef. Typically, a steak is classified as “tough,” if the SSF value is greater than 25 kg.

X-ray Data Analysis

By rearranging Equation 9.1, the attenuation coefficient of a sample at a given voltage can be calculated as:

$$\mu_m = -\frac{1}{\rho t} \ln\left(\frac{I}{I_0}\right) \quad (9.2)$$

For each image, the attenuation coefficient was calculated pixel-by-pixel. The transmitted image, I , was obtained by placing the beef steak between the X-ray source and camera. The incident X-ray intensity or blank image, I_0 , was obtained by acquiring an image without an object between the source and the camera. Both I and I_0 must be acquired at same voltage, current, and integration time. Integration time was set at 460 ms for all images in this study. When a sample was placed between the source and camera, high voltage and current (for instance, 45 kV and 1 mA) were required in order to acquire a quality image. Unfortunately, high voltage and current saturated the photodiode array when there was no material between the source and camera. Under conditions in which the photodiode array did not saturate for blank images (for instance, 20 kV and 1 mA), the transmitted image was too dark. At low voltages, the X-rays were completely attenuated by the sample.

The system was calibrated, and a regression model was developed to estimate I_0 on the basis of input voltage and current:

$$I_0 = A + B(\text{current}) + C(\text{voltage} * \text{current}) \quad (9.3)$$

A , B , and C are regression coefficients. These coefficients were calculated separately for every pixel. Details of calibration are given in Appendix A. Equation 9.2 requires entry of the thickness of the beef steak sample at every pixel. It is almost impossible to measure the thickness of steak accurately, because of pliable muscle surface and small pixel size. Accuracy of thickness measurement is critical for calculation of attenuation coefficient. In this study, we calculated an R-value (another material property related to X-ray attenuation), that does not require thickness measurements.

R-value is defined as the ratio of the attenuation coefficient at low-energy level to the attenuation coefficient at high-energy level (Buzzell and Pintauro, 2004). As the energy of X-rays increases, the attenuation coefficient decreases. However, the rate at which the attenuation coefficient decreases with energy varies from material to material. This material property has been successfully used in dual-energy X-ray absorption techniques for determining biological body composition. The R-value of human muscle is linearly related to the fat content of the muscle (Buzzell and Pintauro, 2004). Mitchell et al. (2003) reported the R-values of pork lean and fat at 38 and 70 keV as 1.4 and 1.2, respectively. The R-value is calculated by:

$$R = \frac{\mu_m^L}{\mu_m^H} = \frac{\ln\left(\frac{I^L}{I_o^L}\right)}{\ln\left(\frac{I^H}{I_o^H}\right)} \quad (9.4)$$

where superscripts ' L ' and ' H ' stand for low- and high-energy levels, respectively.

Equation 9.4 does not require the calculation of thickness. In this study, R-values were calculated (pixel-by-pixel) at various energy combinations (L/H : 30/35, 30/40, 30/45, 35/40, 35/45, 40/45). However, only a single tenderness reference score was

available for the whole steak. Therefore, the mean and standard deviation of these R-value images were calculated and used to predict beef tenderness. In total, 12 features were extracted for each sample.

Results & Discussion

An X-ray image of a beef steak is shown in Figure 9.2. Marbling features are evident in the image. Figure 9.3 shows an R-value image of the same steak. The R-value image looks like a random noise image void of features.

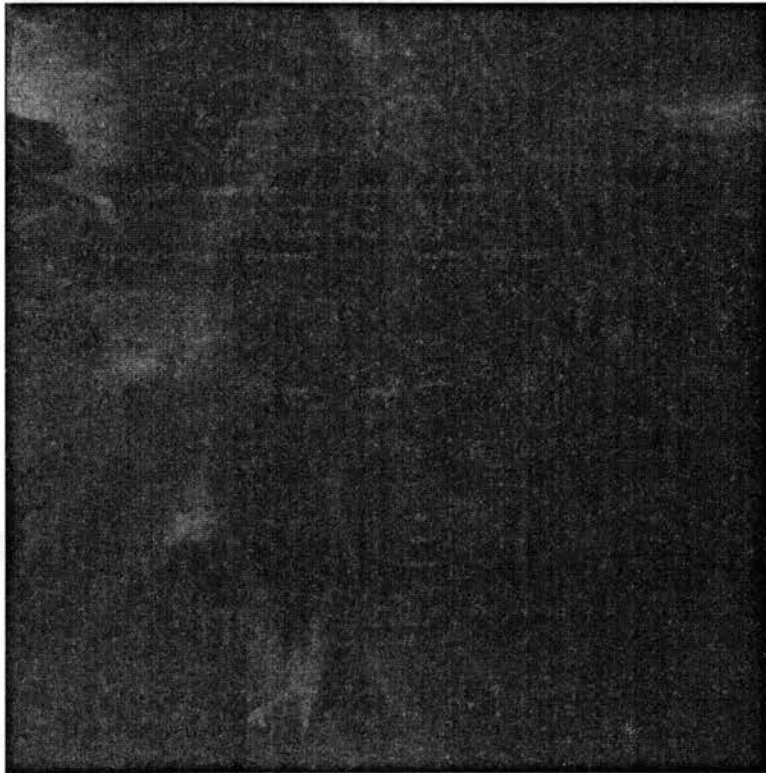


Figure 9.2. X-ray image of a beef steak.

Linear correlation coefficients between the features and slice shear force ranged from 0.02 to 0.09. None of the correlations was significantly different from zero ($\alpha =$

0.05). These results indicate that X-ray attenuation properties of beef are not related to tenderness.

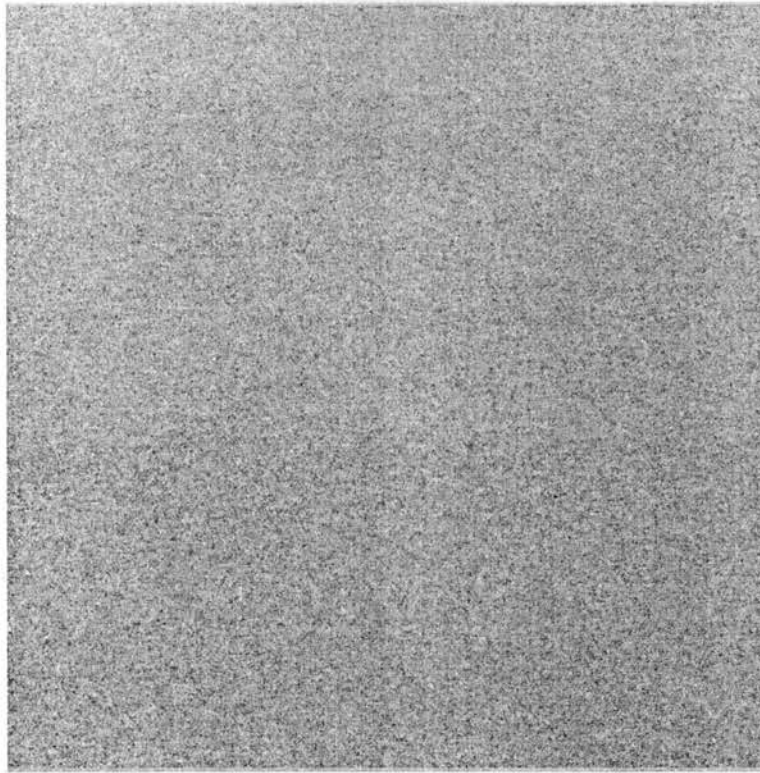


Figure 9.3. R-value image of a beef steak.

In order to detect any weak relationship between attenuation properties and SSF, 10 extreme tender samples ($SSF \leq 11.83$ kg) and 10 extreme tough samples ($SSF \geq 27.61$ kg) were selected. The t-test for independent samples was conducted to detect any significant differences in any of the features between the two extreme categories ($\alpha = 0.05$). There were no significant differences between two categories for any of the 12 features. This result indicates that there was not even a weak relationship between attenuation properties of beef and tenderness.

The X-ray mass attenuation coefficient is affected by molecular composition, but not by chemical structure (Desrosier, 1960). In fact, the mass attenuation coefficient of a

compound can be determined by knowing the atomic weight and weight fraction of all constituents (Berger et al., 1999; Hubbell and Seltzer, 1996). For example, polyethylene has a molecular composition of $[C_2H_4]_n$. By adjusting the processing conditions, low-density polyethylene (LDPE) or high-density polyethylene (HDPE) can be selectively manufactured. LDPE and HDPE have the same molecular composition, but different polymer branching. Densities are very different. Mass attenuation coefficients are identical.

The molecular composition of beef does not change during the aging process. That is, the amounts of carbon, oxygen, hydrogen, and other elements present in the beef muscle immediately after slaughter are the same after aging. Thus, R-value images obtained immediately after slaughter and after aging should be similar. However, enzymatic activity during aging is known to produce chemical changes that help tenderize the beef. Thus, X-ray attenuation properties cannot be used to predict beef tenderness.

Conclusions

A digital soft X-ray imaging system was developed and calibrated. X-ray images of beef steaks were acquired at various energy levels. R-value (ratio of attenuation coefficient at low energy level to that at high energy) images were calculated. Mean and standard deviation of R-value images, obtained at various energy level combinations, were used to predict slice shear-force tenderness values. No suitable model could be developed to predict tenderness. X-ray attenuation properties of beef cannot be used to predict tenderness, because those properties are not affected by chemical changes. This study

does not support the hypothesis that X-ray attenuation properties are related to tenderness.

References

- Berger, M. J., J. H. Hubbell, S. M. Seltzer, J. S. Coursey, and D. S. Zucker. 1999. XCOM: Photon Cross Sections Database. NIST Standard Reference Database 8 (XGAM). Gaithersburg, MD.: National Institute of Standards and Technology. Available at: <http://physics.nist.gov/PhysRefData/Xcom/Text/XCOM.html>. Accessed on 8 March 2004.
- Brienne, J. P., C. Denoyelle, H. Baussart, and J. D. Daudin. 2001. Assessment of meat fat content using dual energy X-ray absorption. *Meat Science* 57(3): 235-244.
- Buzzell, P., and S. Pintauro. 2004. Dual energy X-ray absorbitometry. Burlington, Vermont: Department of Nutrition and Food Sciences, University of Vermont. Available at: <http://nutrition.uvm.edu/bodycomp/dexa/>. Accessed on 10 February 2004.
- Curry, T. S., J. E. Dowdey, and R. C. Murry. 1990. *Christensen's Physics of Diagnostic Radiology*. 4th ed. Media, PA.: William & Wilkins.
- Desrosier, N. W. 1960. *Radiation Technology in Food, Agriculture, and Biology*. Westport, CT.: AVI Publishing Company.

- Hubbell, J. H., and S. M. Seltzer. 1996. NISTIR 5632: Tables of X-Ray Mass Attenuation Coefficients and Mass Energy-Absorption Coefficients from 1 keV to 20 MeV for Elements Z = 1 to 92 and 48 Additional Substances of Dosimetric Interest. Gaithersburg, MD.: National Institute of Standards and Technology. Available at: <http://www.physics.nist.gov/PhysRefData/XrayMassCoef/cover.html>. Accessed on 24 July 2003.
- Mitchell, A. D., M. B. Solomon, and T. S. Rumsey. 1997. Composition analysis of beef rib sections by dual-energy X-ray absorptiometry. *Meat Science* 47(1-2): 115-124.
- Mitchell, A. D., A. M. Scholz, and J. M. Conway. 1998a. Body composition analysis of small pigs by dual-energy x-ray absorptiometry. *Journal of Animal Science* 76(9): 2392-2398.
- Mitchell, A. D., A. M. Scholz, and J. M. Conway. 1998b. Body composition analysis of pigs from 5 to 97 kg by dual energy X-ray absorptiometry. *Applied Radiation Isotopes*, 49: 521-523.
- Mitchell, A. D., A. M. Scholz, V. G. Pursel, and C. M. Evoke-Clover. 1998c. Composition analysis of pork carcasses by dual-energy x-ray absorptiometry. *Journal of Animal Science* 76(8): 2104-2114.
- Mitchell, A. D., A. M. Scholz, and V. G. Pursel. 2001. Total body and regional measurements of bone mineral content and bone mineral density in pigs by dual energy X-ray absorptiometry. *Journal of Animal Science* 79(10): 2594-2604.
- Mitchell, A. D., A. M. Scholz, and V. G. Pursel. 2003. Prediction of pork carcass composition based on cross-sectional region analysis of dual energy X-ray absorptiometry (DXA) scans. *Meat Science* 63(2): 265-271.

Murray, B. C. 2001. A method for the non-invasive measurement of properties of meat.

International Publication No. WO 01/96844.

Shackelford, S. D., T. L. Wheeler, and M. Koohmaraie. 1999. Evaluation of slice shear force as an objective method of assessing beef longissimus tenderness. *Journal of Animal Science* 77(10): 2693-2699.

CHAPTER X

CONCLUSIONS

Summary

The U.S. beef industry relies primarily on the USDA quality grading system to classify carcasses on the basis of expected meat palatability. Because grades are assigned by human inspection, this system is subjective. In order to achieve consistent and accurate evaluation, an instrumented, objective grading system is desired.

Responding to industry needs, a computer vision system was developed to acquire and process images of fresh beef. Segmentation of the longissimus dorsi (l.d.) muscle is the first step in computer vision quality grading, because USDA inspectors assign quality grades by visual appraisal of the l.d. muscle. The l.d. muscle must be separated from the intermuscular fat and extraneous tissue. Small errors in segmentation can lead to incorrect interpretation of intermuscular fat as marbling and result in considerable error in determining marbling score and quality grade. In processing beef images, simple thresholding is typically used to separate fat and lean. In this study, a fuzzy c-means clustering (FCM) algorithm was used to classify image pixels into fat or lean. FCM is more robust and less affected by variations in illumination level and overall color of the meat.

After separating the fat, the most direct procedure for removing extraneous tissue is by morphological operations. The size and shape of the morphological mask and

number of iterations are critical factors for optimal results. In this study, the number of iterations was automatically adjusted from feedback based on the shape of the ribeye, while the size of the mask was fixed. Using convex hull fitting procedures, the number of iterations was automatically optimized for each sample to achieve the most compact l.d. muscle. Implementing morphological operations on down-sampled images and extending the final segmentation results to original resolution lowered classification accuracy minimally (0.04%), but reduced computation time by 40%. Classification error and average error pixel distance of segmentation were 1.97% and 4.4 pixels, respectively. In comparison, classification error and average error pixel distance by expert manual segmentation (indicating imprecision or repeatability error) were 1.19% and 2.3 pixels, respectively. Classification errors were similar to the imprecision of the expert reference, indicating strong performance of our adaptive segmentation algorithm.

After segmenting the *longissimus dorsi* (l.d.) muscle, marbling and color features were extracted to build regression models for predicting marbling and color scores. Quality grade was predicted using an additional regression model incorporating both marbling and color features. The R^2 values of prediction for color score, marbling score, and quality grade were 0.86, 0.64, and 0.76, respectively.

One of the major challenges in developing instrumentation for beef quality grading is that the reference, manual grading, is not precise. For example, two manual graders may not assign the same quality grade and marbling score to the same carcass. During model development, the computer vision system is referenced to the grades assigned by manual graders. The model is compromised by the imprecision of the

reference. Thus, one can not expect the error of the computer vision to be less than the imprecision of the reference.

In order to justify replacement of the current manual grading system by the computer vision system, it is the responsibility of the developer to prove that the instrument predictions are indeed statistically *equivalent* to the scores assigned by the manual graders. Note that “significantly not different” is not the same as “significantly equivalent.” This situation is similar to the bioequivalence problem commonly encountered by the pharmaceutical industry. In this study, equivalence was declared when the computer vision error was less than or equal to the imprecision of the reference. Based on this analysis, grades predicted by the computer vision system were significantly equivalent to the grades assigned by expert graders at $\alpha = 0.05$.

The USDA quality grading system is used as a tool to classify carcasses into groups on the basis of expected meat palatability. Thus, it is an indirect measure of meat palatability. Tenderness is a major palatability factor. USDA quality grading relies heavily on abundance and distribution of marbling in assigning quality grades. Various studies, however, report that marbling shows only a weak relationship with palatability traits. Tough carcasses are found at all levels of marbling. Thus, the current quality grading system does not satisfactorily sort carcasses on the basis of tenderness. Consequently, the beef industry has been looking for a nondestructive carcass evaluation technique to predict aged, cooked-beef tenderness for online.

Developing an instrument to predict beef tenderness continues to be a challenge because of broad variability involved in aging and cooking. To enhance tenderness, it is standard industry practice to age beef for at least 14 days. In general, aging improves

tenderness. However, the amount of tenderization varies considerably from carcass-to-carcass. The beef industry targets an instrument that collects readings on a carcass at 2-3 days postmortem and predicts 14-day aged tenderness. The instrument should be able to account for the range of the changes encountered during the aging process. Tenderness is also a property of the cooked product. The way in which the beef is cooked also affects tenderness palatability. Even though shear-force tenderness scores are collected from samples cooked under controlled conditions, the cooking equipment used is still not uniform. Predicting a property of cooked meat from fresh meat poses a major challenge.

As ribeye images were acquired for predicting USDA quality grade, the natural procedural progression was to extract textural features from those images to predict beef tenderness. A close-up image (resolution: 480 x 640; size: 32 x 43 mm) of a central ribeye area provided more textural information than a full-scale image of a ribeye (resolution: 480 x 640; size: 134 x 179 mm). In addition, the l.d. muscle segmentation operation was eliminated in the close-up image, saving significant computation time. Images were transformed to CIE LAB color space. Second-order statistical textural features were extracted in the spatial domain using gray-level distance histograms. The R^2 value of the regression model for prediction of Warner-Bratzler shear-force was 0.72. Classification into two tenderness categories was achieved with 92% accuracy.

Encouraged by the above results, further studies were conducted to improve performance by extracting textural features in joint spatial-frequency domains. A Gabor filter was used to extract textural energy distribution in four scales (coarser to finer texture) and six orientations (directionality of texture). A canonical discriminant model classified carcasses into three tenderness categories, with 80% accuracy. When the

model classified two tenderness categories, accuracy improved to 92%. Thus, performance of the Gabor filter was similar to that of the gray-level distance histograms.

A four-scale and six-orientation Gabor filter decomposes an image into 24 images with distinct frequency bands and orientations. As a result, memory requirements are high. Wavelets also decompose an image into several subimages of distinct frequency band and of three orientations. Down-sampling is performed at each level of wavelet decomposition to remove redundant information. Unlike the Gabor filter, memory requirements do not increase for wavelet analysis. In our study, wavelet textural features predicted shear-force values with a low correlation coefficient of 0.53. A tenderness certification method was used to evaluate the ability of the system to sort carcasses. At all tenderness certification levels, significant differences were observed between the mean observed shear-force values of two certified groups. This result indicated that the wavelet textural features were able to successfully sort the carcasses on the basis of tenderness. The performance of wavelet textural features cannot be directly compared to that of Gabor and statistical textural features, because performance evaluation methods were different. In summary, the Gabor filter and wavelets did not provide a significant advantage over gray-level distance histograms.

Computer vision in the visible region of the spectrum simulates human vision. Tender beef has fine muscle fibers, whereas tough beef has visibly coarser muscle fibers. Fresh beef images in the visible region do not contain explicit chemical/biochemical information. Light reflected from the beef sample in the near-infrared (NIR) region of the spectrum does contain information about chemical properties. Because tenderness is thought to be related to meat chemical properties affected by aging, NIR offers higher

potential for information related to aging, in addition to the current state of tenderness. Thus, NIR spectroscopy was evaluated for predicting beef tenderness.

In this dissertation, NIR was more rigorously tested than computer vision. For computer vision evaluation, steaks were imaged in the laboratory (offline). Evaluation of the NIR system was conducted in packing plants. Spectral scans of the ribeye on hanging carcasses were acquired online. Further, performance of the NIR system was validated by an outside party. For validation, carcasses (n=200) were scanned online in the packing plant and the model developed was used to predict tenderness scores. After scanning, a steak was cut from each carcass, aged for 14 days, and frozen. Frozen steaks, along with predicted tenderness categories, were forwarded to the U.S. Meat Animal Research Center (MARC) in Clay Center, Nebraska for validation. MARC cooked the samples and collected slice-shear force measurements in their laboratory. MARC cooking equipment and cooking time varied from procedures used at Oklahoma State University for model development. Because of these differences, validation results were slightly below our calibration results. Nevertheless, significant differences between the slice-shear force values of “certified tender” and “not certified tender” were observed at or below 70% tenderness certification levels. The system successfully sorted “tough” from “tender” carcasses at or below 70% tenderness certification levels. Practical implications for the beef industry are that, if 30% of carcasses sorted “tough” by this system are removed, the remaining 70% can be sold as “guaranteed tender” to premium markets such as restaurants.

For an earlier food quality inspection project, a soft X-ray digital imaging system was developed at OSU Biosystems and Agricultural Engineering. In this dissertation, a

feasibility study was conducted to determine whether X-ray attenuation properties of beef could be used to predict tenderness. The system was calibrated and a regression model was developed to predict incident X-ray intensity, based on input voltage and current. R-values (ratio of attenuation coefficient at low-energy to that at high-energy) were calculated at several energy combinations. No correlation was found between X-ray attenuation and tenderness. No suitable model could be developed to predict beef tenderness. This study does not support the hypothesis that attenuation properties are related to tenderness.

Suggestions for Future Research

This dissertation evaluated three nondestructive technologies and concluded that two, computer vision and NIR spectroscopy, showed promise for predicting aged, cooked-beef tenderness. The NIR system used in this study collected spectral scans from a 1-cm diameter area on the ribeye surface. A foreoptic can be fitted to the end of fiberoptic probe to widen the field-of-view. Implementation would allow collecting spectral scans from a larger area on ribeye surface. Fresh beef reflects less than 40% of incident light in the visible-NIR region. In this study, a white reference plate was used to optimize integration time. Instead of the white reference, a 50% gray plate could be used to optimize integration time. This change would double the integration time, improving signal strength to increase information gathered.

In this study, orthogonal wavelet transform was used, which produces rotationally-variant features. Rotation of the image would change these wavelet features. To reduce this effect, details at three orientations were added. This procedure did not completely eliminate the effect of rotation on features. There are non-orthogonal

(redundant) wavelet transforms such as steerable wavelet pyramids that produce rotationally-invariant and translationally-invariant features. Non-orthogonal wavelet transforms should be evaluated for textural analysis of beef images. Our computer vision system was developed to predict beef tenderness offline. The system should be modified for online, packing-plant implementation. For online operation, the camera should be fitted to a frame with a shrouded lighting system. Calibration procedures for illumination and color would need to be developed and implemented. Performance evaluation in a packing plant environment would be required.

Performance of both NIR and computer vision systems can be improved by incorporating supplemental factors such as sex, age, and breed of the animal, carcass weight, geographical region of origin, type of feed, etc. This additional information will be more easily accessible in future because of more rigorous animal tracking spurred by the recent threat of mad cow disease.

NIR can probe chemical characteristics of a food material on the basis of intramolecular vibrational frequencies, but define the internal structure. X-rays can provide information on internal structure of materials, but cannot provide details on chemical characteristics. Magnetic resonance can provide details on both chemical characteristics and the internal structure of materials. Magnetic resonance imaging (MRI) has been successful for investigating composition analysis and water holding capacity of meat. Only two studies have shown that MRI images can sense tenderness-related connective tissue. No work has been done using MRI to predict beef tenderness. This topic would be of interest for future investigation.

We found that both image texture and spectral reflectance were useful in predicting tenderness. The next logical step is to evaluate hyperspectral imaging, which is an extension of both computer vision and NIR spectroscopy. A typical RGB (red-green-blue) camera captures image in three broadband spectral regions. A spectrometer captures spectral content of a small window, without spatial resolution. In contrast, a hyperspectral camera collects spectral contents of every pixel of an image, with high spectral resolution. This technology produces a 3-D data cube (x, y, λ) . Thus, the hyperspectral camera can be considered as an extension of an RGB camera with higher resolution on the spectral axis (λ) .

One way to capture hyperspectral data is by use of a non-scanning, computed tomography imaging spectrometer (CTIS). The Optics Science Center, University of Arizona is leading research in the CTIS area. They designed a special grating to project the data cube on a focal-plane 2-D array of detectors in selected directions. The 3-D data cube was then reconstructed from those 2-D image projections using computed tomography techniques. One advantage is that this method is non-scanning, and thus the capture of hyperspectral data is rapid. The disadvantage is that a trade-off exists between spectral and spatial resolution. CTIS is still in the research stage and may be useful in future for collecting hyperspectral data.

The most common method of collecting hyperspectral imaging data is by using liquid crystal tunable imaging filters. These filters are commercially available and can be fitted to the lens of a broad-spectrum camera. These filters pass a specific spectral band of reflected light for imaging. By tuning these filters, the band-pass region can be shifted throughout the designed spectral range. Switching speed between bands is currently

about 150 ms. If the integration time is set at 150 ms, roughly three spectral bands can be acquired per second. Thus, this method is slower than CTIS. Unlike CTIS, spatial resolution depends only on the camera and is not compromised by spectral resolution. In application, textural features could be extracted from each spectral band. Drawing from preliminary studies, the most productive bands would be selected. Only the selected band-pass filters would be installed, considerably reducing acquisition time.

The effectiveness of computer vision and NIR systems in predicting tenderness of other middle and end muscle cuts such as *triceps brachii*, *rectus femoris*, *gluteus medius*, *semimembranosus*, and *semitendinosus*, should be evaluated. An experiment can be designed to determine the relationship between the shear-force values of l.d. muscle and the shear-force values of other muscle cuts. If a strong relationship exists, the instrument can collect readings from l.d. muscle and use them to predict tenderness ratings of other muscles from the same carcass.

In this dissertation, shear-force values were used as the tenderness reference. Trained sensory panels can be used to assign scores for tenderness, juiciness, flavor, and overall palatability. Collecting sensory panel scores is more time-consuming and labor-intensive than measuring shear-force values. However, sensory scores are more representative of consumer perception of beef palatability. Shear-force values and sensory scores should be combined as the palatability reference for future work.

The computer vision and NIR system can be used to classify the carcasses into three tenderness categories. A study can be conducted to determine consumer perception of these categories. Two steaks from each category ($n = 6$) can be delivered to each household. Steaks can be color-coded by category so that the consumer does not know

the tenderness class. After eating the steaks, consumers would be asked to complete a survey. Based on consumer response, effectiveness of these nondestructive methods can be evaluated on the basis of consumer perceptions.

COMPLETE REFERENCE LIST

- Allen, P., and N. Finnerty. 2000. Objective beef classification: A report of a trial of three VIA classification systems. Dublin, Ireland: The National Food Centre, Teagasc. Available at: <http://www.teagasc.ie/publications/beefgrading/beefgrading.pdf>. Accessed on 16 January 2004.
- AMSA. 1995. *Research guidelines for cookery, sensory evaluation, and instrumental tenderness measurements of fresh meat*. Chicago, IL.: National Live Stock & Meat Board, American Meat Science Association.
- Antequera, T., E. Muriel, P. G. Rodríguez, E. Cernadas, and J. Ruiz. 2003. Magnetic resonance imaging as a predictive tool for sensory characteristics and intramuscular fat content of dry-cured loin. *Journal of the Science of Food and Agriculture* 83(4): 268-274.
- Baraldi, A., and F. Parmiggiani. 1995. An investigation of the textural characteristics associated with gray level cooccurrence matrix statistical parameters. *IEEE Transactions on Geoscience and Remote Sensing* 33(2): 293-304.
- Barni, M., V. Cappellini, and A. Mecocci. 1997. Colour-based detection of defects on chicken meat. *Image Vision Computing* 15: 549-556.
- Belk, K. E., J. A. Scanga, J. D. Tatum, J. W. Wise, and G. C. Smith. 1998. Simulated instrument augmentation of USDA yield grade application to beef carcasses. *Journal of Animal Science* 76(2): 522-527.

- Belk, K. E., J. D. Tatum, G. C. Smith, M. Goldberg, A. M. Wyle, and R. C. Cannell. 2001. Meat imaging system for palatability yield prediction. U.S. Patent No. 5793893.
- Belk, K. E., J. D. Tatum, G. Dolezal, B. Morgan, and G. C. Smith. 1996. Meat composition measurement: status of applied research on instrument assessment of composition since completion of the 1994 National Beef Instrument Assessment Planning Symposium. In *Reciprocal Meat Conference Proceedings*, 49: 172-174.
- Belk, K. E., M. H. George, J. D. Tatum, G. G. Hilton, R. K. Miller, M. Koohmaraie, J. O. Reagan, and G. C. Smith. 2001. Evaluation of the Tendertec beef grading instrument to predict the tenderness of steaks from beef carcasses. *Journal of Animal Science* 79(3): 688-697.
- Benet, L. Z. 1999. Understanding bioequivalence testing. *Transplantation Proceedings* 31(Suppl 3A): 7S-9S.
- Benn, A., D. Barrett-Lennard, and P. J. Hay. 1998. Image analysis for meat. U.S. Patent No. 5793879.
- Berger, M. J., J. H. Hubbell, S. M. Seltzer, J. S. Coursey, and D. S. Zucker. 1999. XCOM: Photon Cross Sections Database. NIST Standard Reference Database 8 (XGAM). Gaithersburg, MD.: National Institute of Standards and Technology. Available at: <http://physics.nist.gov/PhysRefData/Xcom/Text/XCOM.html>. Accessed on 8 March 2004.
- Bertram, H. C., A. K. Whittaker, W. R. Shorthose, H. J. Andersen, and A. H. Karlsson. 2004. Water characteristics in cooked beef as influenced by ageing and high-pressure treatment--an NMR micro imaging study. *Meat Science* 66(2): 301-306.

- Bezdek, C. J. 1981. *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York, NY.: Plenum Press.
- Bezdek, C. J., J. Keller, R. Krishnapuram, and N. R. Pal. 1999. *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing*. Boston, MA.: Kluwer Academic Publishers.
- Biju, N. 1998. Beef quality grading with color video image analysis. MS thesis. Stillwater, Oklahoma: Oklahoma State University, Department of Biosystems and Agricultural Engineering.
- Boleman, S. J., S. L. Boleman, R. K. Miller, J. F. Taylor, H. R. Cross, T. L. Wheeler, M. Koohmaraie, S. D. Shackelford, M. F. Miller, R. L. West, D. D. Johnson, and J. W. Savell. 1997. Consumer evaluation of beef of known categories of tenderness. *Journal of Animal Science* 75(6): 1521-1524.
- Boleman, S. L., S. J. Boleman, W. W. Morgan, D. S. Hale, D. B. Griffin, J. W. Savell, R. P. Ames, M. T. Smith, J. D. Tatum, T. G. Field, G. C. Smith, B. A. Gardner, J. B. Morgan, S. L. Northcutt, H. G. Dolezal, D. R. Gill, and F. K. Ray. 1998. National Beef Quality Audit-1995: survey of producer-related defects and carcass quality and quantity attributes. *Journal of Animal Science* 76(1): 96-103.
- Bonny, J.-M., W. Laurent, and J.-P. Renou. 2000a. Detection of susceptibility effects using simultaneous T2* and magnetic field mapping. *Magnetic Resonance Imaging* 18(9): 1125-1128.

- Bonny, J.-M., W. Laurent, R. Labas, R. Taylor, P. Berge, and J.-P. Renou. 2000b. Magnetic resonance imaging of connective tissue: a non-destructive method for characterising muscle structure. *Journal of the Science of Food and Agriculture* 81(3): 337 - 341.
- Borggaard, C., N. T. Madsen, and H. H. Thodberg. 1996. In-line image analysis in the slaughter industry, illustrated by Beef Carcass Classification. *Meat Science* 43(Supplement 1): 151-163.
- Brienne, J. P., C. Denoyelle, H. Baussart, and J. D. Daudin. 2001. Assessment of meat fat content using dual energy X-ray absorption. *Meat Science* 57(3): 235-244.
- Brooks, J. C., J. B. Belew, D. B. Griffin, B. L. Gwartney, D. S. Hale, W. R. Henning, D. D. Johnson, J. B. Morgan, F. C. Parrish, Jr, J. O. Reagan, and J. W. Savell. 2000. National Beef Tenderness Survey-1998. *Journal of Animal Science* 78(7): 1852-1860.
- Butler, J., J. Marchello, E. E. Ray, and R. Shaw. 1980. Standards for Beef Color. Iowa State Extension Publication: AS-515. Ames, IA.: Iowa State University.
- Buzzell, P., and S. Pintauro. 2004. Dual energy X-ray absorbitometry. Burlington, Vermont: Department of Nutrition and Food Sciences, University of Vermont. Available at: <http://nutrition.uvm.edu/bodycomp/dexa/>. Accessed on 10 February 2004.
- Byrne, C. E., G. Downey, D. J. Troy, and D. J. Buckley. 1998. Non-destructive prediction of selected quality attributes of beef by near-infrared reflectance spectroscopy between 750 and 1098 nm. *Meat Science* 49(4): 399-409.
- CAMO. 1998. The Unscrambler User Guide. Ver. Corvallis, OR.: Camo, Inc.

- Cannell, R. C., J. D. Tatum, K. E. Belk, J. W. Wise, R. P. Clayton, and G. C. Smith. 1999. Dual-component video image analysis system (VIASCAN) as a predictor of beef carcass red meat yield percentage and for augmenting application of USDA yield grades. *Journal of Animal Science* 77(11): 2942-2950.
- Connolly, C., and T. Fleiss. 1997. A study of efficiency and accuracy in the transformation from RGB to CIELAB color space. *Image Processing, IEEE Transactions on* 6(7): 1046-1048.
- Cormen, T. H., R. L. Leiserson, and R. L. Rivest. 1989. *Introduction to Algorithms*. New York, NY.: McGraw-Hill.
- Cross, H. R., and A. D. Whittaker. 1992. The role of instrument grading in a beef value-based marketing system. *Journal of Animal Science* 70: 984-989.
- Cross, H. R., and K. E. Belk. 1994. Objective measurements of carcass and meat quality. *Meat Science* 36(1-2): 191-202.
- Cross, H. R., D. A. Gilliland, P. R. Durland, and S. C. Seideman. 1983. Beef carcass evaluation by use of a video image analysis system. *Journal of Animal Science* 57: 908-917.
- 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. Final report to office of inspector general, USDA.
- Curry, T. S., J. E. Dowdey, and R. C. Murry. 1990. *Christensen's Physics of Diagnostic Radiology*. 4th ed. Media, PA.: William & Wilkins.

- Dannenbergh, O., H. Dette, and A. Munk. 1994. An extension of Welch's approximate t-solution to comparative bioequivalence trials. *Biometrika* 81(1): 91-101.
- Denoyelle, C., and F. Berny. 1999. Objective measurement of veal color for classification purposes. *Meat Science* 53: 203-209.
- Desrosier, N. W. 1960. *Radiation Technology in Food, Agriculture, and Biology*. Westport, CT.: AVI Publishing Company.
- Esbenson, K. H. 2001. *Multivariate data analysis - in practice. An introduction to multivariate data analysis and experimental design*. 5th ed. Corvallis, OR.: Camo, Inc.
- Foley, J. D., A. V. Dam, S. K. Feiner, and J. F. Hughes. 1995. *Computer Graphics*. 2nd ed. Boston, MA.: Addison Wesley.
- Gao, X., J. Tan, and D. Gerrard. 1995. Image segmentation in 3-dimensional color space. ASAE Paper No. 95-3607. St. Joseph, Mich.: ASAE.
- George, M. H., J. D. Tatum, H. G. Dolezal, J. B. Morgan, J. W. Wise, C. R. Calkins, T. Gordon, J. O. Reagan, and G. C. Smith. 1997. Comparison of USDA quality grade with tendertec for the assessment of beef palatability. *Journal of Animal Science* 75(6): 1538-1546.
- Gerrard, D. E., X. Gao, and J. Tan. 1996. Beef marbling and color score determination by image processing. *Journal of Food Science* 61(1): 145-148.
- Goering, K. 1999. Instrument grading: A different sort of measurement. *National Cattlemen*, 14(1): 54-56.
- Goldenberg, A. A., and Z. Lu. 1997. Automation of meat pork grading process. *Computers and Electronics in Agriculture* 16(2): 125-135.

- Gonzalez, R. C., and R. E. Woods. 1992. *Digital Image Processing*. Reading, M.A.: Addison-Wesley Publishing Company.
- Hammer, B. E. 1998. Industrial applications of nuclear magnetic resonance. *Sensor Review* 18(4): 245-251.
- Haralick, R. M., K. Shanmugam, and I. Dinstein. 1973. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics* SMC-3(6): 610-621.
- Hatem, I., and J. Tan. 2003. Cartilage and bone segmentation in vertebra images. *Transactions of the ASAE* 46(5): 1429-1434.
- Hatem, I., J. Tan, and D. E. Gerrard. 2003. Determination of animal skeletal maturity by image processing. *Meat Science* 65(3): 999-1004.
- Hildrum, K. I., B. N. Nilsen, M. Mielnik, and T. Naes. 1994. Prediction of sensory characteristics of beef by near-infrared spectroscopy. *Meat Science* 38(1): 67-80.
- Hildrum, K. I., T. Isaksson, T. Næs, B. N. Nilsen, M. Rødbotten, and P. Lea. 1995. Near infrared reflectance spectroscopy in the prediction of sensory properties of beef. *Journal of Near Infrared Spectroscopy* 3: 81-87.
- Holder, D. J., and F. Hsuan. 1993. Moment based criteria for determining bioequivalence. *Biometrika* 80(4): 835-846.
- Hruschka, W. R. 2001. Chapter 3: Data Analysis: Wavelength selection methods. In *Near-Infrared Technology in the Agricultural and Food Industries.*, P. Williams and K. Norris., eds. St. Paul, MN.: American Society of Cereal Chemist, Inc.
- Hsu, J. C., J. T. G. Hwang, H. Liu, and S. J. Ruberg. 1994. Confidence intervals associated with tests for bioequivalence. *Biometrika* 81(1): 103-114.

- Huang, Y., R. E. Lacey, L. L. Moore, R. K. Miller, A. D. Whittaker, and J. Ophir. 1997. Wavelet textural features from ultrasonic elastograms for meat quality prediction. *Transactions of the ASAE* 40(6): 1741-1748.
- Hubbell, J. H., and S. M. Seltzer. 1996. NISTIR 5632: Tables of X-Ray Mass Attenuation Coefficients and Mass Energy-Absorption Coefficients from 1 keV to 20 MeV for Elements Z = 1 to 92 and 48 Additional Substances of Dosimetric Interest. Gaithersburg, MD.: National Institute of Standards and Technology. Available at: <http://www.physics.nist.gov/PhysRefData/XrayMassCoef/cover.html>. Accessed on 24 July 2003.
- IEEE Standard 610-4. 1990. *IEEE Standard Glossary of Image Processing and Pattern Recognition Terminology*. New York, NY.: IEEE Press.
- Ishmael, W. 2000. Value Vision. Available at: http://beef-mag.com/ar/beef_value_vision/index.htm. Accessed on 6 February 2004.
- Jain, A. K. 1989. *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ.: Prentice Hall.
- Jeremiah, L. E. 1996. The influence of subcutaneous fat thickness and marbling on beef: Palatability and consumer acceptability. *Food Research International* 29(5-6): 513-520.
- Jeremiah, L. E., and D. M. Phillips. 2000. Evaluation of a probe for predicting beef tenderness. *Meat Science* 55(4): 493-502.
- Jeyamkondan, S., and G. A. Kranzler. 2001. Bioequivalence analysis for beef quality grading with computer vision. In *2001 Oklahoma Section Meeting of the American Society of Agricultural Engineers*. Stillwater, OK.

- Jeyamkondan, S., N. Ray, G. A. Kranzler, and B. Nisha. 2000. Beef quality grading using machine vision. In *Proceedings of SPIE*, 4203: 91-101. J. A. DeShazer and G. E. Meyer, eds. Boston, MA.: SPIE.
- Jeyamkondan, S., N. Ray, G. A. Kranzler, and S. Acton. 2004. Computer vision segmentation of the longissimus dorsi for beef quality grading. *Accepted for publication in the Transactions of the ASAE*.
- Johnson, D. E. 1998. *Applied Multivariate Methods for Data Analysis*. New York, NY.: Duxbury Press.
- Johnson, H. K., and T. R. Dockerty. 1990. *Recommended Procedures for Beef Carcass Evaluation and Carcass Contests*. 3rd ed. Savoy, IL.: American Meat Science Association.
- Kim, N., D. Wilson, G. Rouse, and S. Udpa. 1998. Ultrasonic image texture analysis for characterizing intramuscular fat content of live beef cattle. *Ultrasonic Imaging* 20: 191-205.
- Kinsella, A. 1989. Biostrapping a bioequivalence measure. *Statistician* 38: 175-179.
- Koohmaraie, M. 1995. The biological basis of meat tenderness and potential genetic approaches for its control and prediction. Clay Center, NE.: USDA-ARS, Roman L. Hruska U.S. Meat Animal Research Center. Available at: http://meats.marc.usda.gov/MRU_WWW/ICMST95/ICMST95.html. Accessed on 9 February 2004.

- Koohmaraie, M. 1996. Biochemical factors regulating the toughening and tenderization processes of meat. Clay Center, NE.: USDA-ARS, Roman L. Hruska U.S. Meat Animal Research Center,. Available at:
http://meats.marc.usda.gov/MRU_WWW/ICMST96/ICMST96.html. Accessed on 9 February 2004.
- Koohmaraie, M. 1999. MARC beef classification system: objective evaluation of beef tenderness and cutability. Clay Center, NE.: Roman L. Hruska U. S. Meat Animal Research Center, Agricultural Research Service, United States Department of Agriculture. Available at: http://meats.marc.usda.gov/MRU_WWW/Bulletin.pdf. Accessed on 17 February 2004.
- Kranzler, G. A., and M. P. Rigney. 1989. Machine vision grading of tree seedlings. In *Proceedings of the 11th International Congress in Agricultural Engineering*. 1883-1888. Rotterdam, The Netherlands.
- Leemans, V., H. Magein, and M. F. Destain. 1998. Defects segmentation on 'golden delicious' apples by using colour machine vision. *Computers and Electronics in Agriculture* 20: 117-130.
- Lenhart, D. H., and D. A. Gilliland. 1985. The design and testing of an automated beef grader. ASAE Paper No. 85-3035. St. Joseph, Mich.: ASAE.
- Leroy, B., S. Lambotte, O. Dotreppe, H. Lecocq, L. Istasse, and A. Clinquart. 2003. Prediction of technological and organoleptic properties of beef Longissimus thoracis from near-infrared reflectance and transmission spectra. *Meat Science* 66(1): 45-54.

- Li, J., J. Tan, and P. Shatadal. 1999. Discrimination of beef images by textural features. ASAE Paper No. 99-3158. St. Joseph, Mich.: ASAE.
- Li, J., J. Tan, and P. Shatadal. 2001. Classification of tough and tender beef by image texture analysis. *Meat Science* 57(4): 341-346.
- Li, J., J. Tan, F. A. Martz, and H. Heymann. 1999. Image texture features as indicators of beef tenderness. *Meat Science* 53(1): 17-22.
- Liu, Y., B. G. Lyon, W. R. Windham, C. E. Realini, T. D. D. Pringle, and S. Duckett. 2003. Prediction of color, texture, and sensory characteristics of beef steaks by visible and near infrared reflectance spectroscopy. A feasibility study*. *Meat Science* 65(3): 1107-1115.
- Livens, S., P. Scheunders, G. Van De Wouwer, and D. Van Dyck. 1997. Wavelets for texture analysis, an overview. In *Sixth International Conference on Image Processing and Its Applications*, 2: 581-585.
- Lobenberg, R., and G. L. Amidon. 2000. Modern bioavailability, bioequivalence and biopharmaceutics classification system. New scientific approaches to international regulatory standards. *European Journal of Pharmaceutics and Biopharmaceutics* 50: 3-12.
- Lorenzen, C. L., D. S. Hale, D. B. Griffin, J. W. Savell, K. E. Belk, T. L. Frederick, M. F. Miller, T. H. Montgomery, and G. C. Smith. 1993. National Beef Quality Audit: survey of producer-related defects and carcass quality and quantity attributes. *Journal of Animal Science* 71(6): 1495-1502.
- Lu, J., and J. Tan. 1998. Application of nonlinear transforms to beef image processing. ASAE Paper No. 98-3016. St. Joseph, Mich.: ASAE.

- Lu, W., and J. Tan. 2004. Analysis of image-based measurements and USDA characteristics as predictors of beef lean yield. *Meat Science* 66(2): 483-491.
- Lusk, J. L., J.A. Fox, T.C. Schroeder, J. Mintert, and M. Koohmaraie. 2001. In-store valuation of steak tenderness. *American Journal of Agricultural Economics* 83(3): 539-550.
- Madsen, N. T., and H. H. Thodberg. 1999. Objective beef classification and quality measurements. In *European experts colloquium on grading and classification*. Kulmbach, Germany: European Commission Fair, MACA Meeting.
- Manjunath, B. S., and W. Y. Ma. 1996. Texture features for browsing and retrieval of image data. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18(8): 837-842.
- Matlab. 2001. Image Processing Toolbox for Use with Matlab® : User's Guide. Ver. 3. Natick, MA.: The MathWorks, Inc.
- Mcdonald, T. P., and Y. R. Chen. 1990. Separating connected muscle tissues in images of beef carcass ribeyes. *Transactions of the ASAE* 33(6): 2059-2065.
- McKenna, D. R., D. L. Roebert, P. K. Bates, T. B. Schmidt, D. S. Hale, D. B. Griffin, J. W. Savell, J. C. Brooks, J. B. Morgan, T. H. Montgomery, K. E. Belk, and G. C. Smith. 2002. National Beef Quality Audit-2000: survey of targeted cattle and carcass characteristics related to quality, quantity, and value of fed steers and heifers. *Journal of Animal Science* 80(5): 1212-1222.
- Misiti, M., Y. Misiti, G. Oppenheim, and J. Poggi. 1996. Wavelet Toolbox for use with Matlab - User's Guide. Ver. 1. Natick, MA.: The Mathworks, Inc.

- Mitchell, A. D., A. M. Scholz, V. G. Pursel, and C. M. Evoke-Clover. 1998a. Composition analysis of pork carcasses by dual-energy x-ray absorptiometry. *Journal of Animal Science* 76(8): 2104-2114.
- Mitchell, A. D., A. M. Scholz, and J. M. Conway. 1998b. Body composition analysis of small pigs by dual-energy x-ray absorptiometry. *Journal of Animal Science* 76(9): 2392-2398.
- Mitchell, A. D., A. M. Scholz, and J. M. Conway. 1998c. Body composition analysis of pigs from 5 to 97 kg by dual energy X-ray absorptiometry. *Applied Radiation Isotopes*, 49: 521-523.
- Mitchell, A. D., A. M. Scholz, P. C. Wange, and H. Song. 2001a. Body composition analysis of the pig by magnetic resonance imaging. *Journal of Animal Science* 79(7): 1800-1813.
- Mitchell, A. D., A. M. Scholz, and V. G. Pursel. 2001b. Total body and regional measurements of bone mineral content and bone mineral density in pigs by dual energy X-ray absorptiometry. *Journal of Animal Science* 79(10): 2594-2604.
- Mitchell, A. D., A. M. Scholz, and V. G. Pursel. 2003. Prediction of pork carcass composition based on cross-sectional region analysis of dual energy X-ray absorptiometry (DXA) scans. *Meat Science* 63(2): 265-271.
- Mitchell, A. D., M. B. Solomon, and T. S. Rumsey. 1997. Composition analysis of beef rib sections by dual-energy X-ray absorptiometry. *Meat Science* 47(1-2): 115-124.
- Mitsumoto, M., S. Maeda, T. Mitsuhashi, and S. Ozawa. 1991. Near-infrared spectroscopy determination of physical and chemical characteristics in beef cuts. *Journal of Food Science* 59(6): 1493-1496.

- Morgan, J. B., J. W. Savell, D. S. Hale, R. K. Miller, D. B. Griffin, H. R. Cross, and S. D. Shackelford. 1991. National beef tenderness survey. *Journal of Animal Science* 69(8): 3274-3283.
- Munk, A. 2000. An unbiased test for the average equivalence problem - the small sample case. *Journal of Statistical Planning Inference* 87: 69-86.
- Murray, B. C. 2001. A method for the non-invasive measurement of properties of meat. International Publication No. WO 01/96844.
- NCBA. 2000. The U.S. beef industry: its impact on the American economy. Fact Sheet. Centennial, CO.: National Cattlemen's Beef Association.
- NCBA. 2001. The U.S. beef industry: its impact on the American economy. Fact Sheet. Centennial, CO.: National Cattlemen's Beef Association. Available at:
http://www.beef.org/dsp/dsp_content.cfm?locationId=710&contentType=1&contentId=252. Accessed on 11 February 2004.
- NCBA. 2002. Beef Update. Meeting Summary. National Beef Instrument Assessment Plan II: Focus on Tenderness. Centennial, CO: National Cattlemen's Beef Association. Available at:
<http://www.beefboard.org/documents/beef%20update.pdf>. Accessed on 9 February 2004.
- NCBA. 2003. Economics of Beef. Beef Bytes. Centennial, CO.: National Cattlemen's Beef Association. Available at:
<http://www.beef.org/documents//Beef%20Bytes%20Economics.pdf>. Accessed on 24 July 2003.

- Newman, P. B. D. 1995. Grading and quality control of meat cuts. Great Britain Patent No. 2285126.
- Nielson, T., N. T. Madsen, and H. H. Thodberg. 1997. Technology transfer in computer vision - Example: beef carcass classification centre BCC-2. In *Technology Transfer in Computer Vision*. Lappeeranta, Finland: ECVnet.
- O"Quigley, J., and C. Baudoin. 1988. General approaches to the problem of bioequivalence. *Statistician* 37: 51-58.
- OBIC. 2003. Oklahoma Beef Statistics. Oklahoma City, OK.: Oklahoma Beef Council. Available at: <http://www.oklabeef.org/Beef%20Statistics.htm>. Accessed on 24 July 2003.
- Oppenheim, A. V., and R. W. Schaffer. 1999. *Discrete-Time Signal Processing*. 2nd ed. New Jersey, NJ.: Prentice Hall.
- Park, B., Y. R. Chen, W. R. Hruschka, S. D. Shackelford, and M. Koohmaraie. 1998. Near-infrared reflectance analysis for predicting beef longissimus tenderness. *Journal of Animal Science* 76(8): 2115-2120.
- Perkins, T., A. Meadows, and B. Hays. 2003. Study Guide for the Ultrasonic evaluation of beef cattle for carcass merit. Ames, IA.: Ultrasound Guidelines Council, Beef Cattle Ultrasound Technician Annual Proficiency Testing and Certification, Iowa State University. Available at: <http://www.aptcbeef.org/UGC%20STUDY%20GUIDE.pdf>. Accessed on 11 February 2004.

- Petersen, F., S. Klastrup, N. T. Madsen, and S. E. Sorensen. 1989. Beef Classification Center. In *Proceedings of the International Congress of Meat Science and Technology*, 35: 49-52.
- Pollikar, R. 2002. The Wavelet Tutorial. Glassboro, NJ.: Department of Electrical and Computer Engineering, Rowan University. Available at:
<http://users.rowan.edu/~polikar/WAVELETS/WTtutorial.html>. Accessed on 21 July 2002.
- Poynton, C. 1999. Frequently asked questions about color. Available at:
<http://www.poynton.com/PDFs/ColorFAQ.pdf>. Accessed on 15 February 2004.
- Renou, J. P., L. Foucat, and J. M. Bonny. 2003. Magnetic resonance imaging studies of water interactions in meat. *Food Chemistry* 82(1): 35-39.
- Rødbotten, M., B. N. Nilsen, and K. I. Hildrum. 2000. Prediction of beef quality attributes from early post mortem near infrared reflectance spectra. *Food Chemistry* 69(4): 427-436.
- Rødbotten, R., B.-H. Mevik, and K. I. Hildrum. 2001. Prediction and classification of tenderness in beef from non-invasive diode array detected NIR spectra. *Journal of Near Infrared Spectroscopy* 9: 199-210.
- Serra, J. 1982. *Image Analysis And Mathematical Morphology*. New York, NY.: Academic Press.
- Shackelford, S. D., T. L. Wheeler, and M. Koohmaraie. 1997. Tenderness classification of beef: I. Evaluation of beef longissimus shear force at 1 or 2 days postmortem as a predictor of aged beef tenderness. *Journal of Animal Science* 75(9): 2417-2422.

- Shackelford, S. D., T. L. Wheeler, and M. Koohmaraie. 1998. Coupling of image analysis and tenderness classification to simultaneously evaluate carcass cutability, longissimus area, subprimal cut weights, and tenderness of beef. *Journal of Animal Science* 76(10): 2631-2640.
- Shackelford, S. D., T. L. Wheeler, and M. Koohmaraie. 1999. Evaluation of slice shear force as an objective method of assessing beef longissimus tenderness. *Journal of Animal Science* 77(10): 2693-2699.
- Shackelford, S. D., T. L. Wheeler, and M. Koohmaraie. 1999. Tenderness classification of beef: II. Design and analysis of a system to measure beef longissimus shear force under commercial processing conditions. *Journal of Animal Science* 77(6): 1474-1481.
- Shackelford, S. D., T. L. Wheeler, and M. Koohmaraie. 2003. On-line prediction of yield grade, longissimus muscle area, preliminary yield grade, adjusted preliminary yield grade, and marbling score using the MARC beef carcass image analysis system. *Journal of Animal Science* 81(1): 150-155.
- Shackelford, S. D., T. L. Wheeler, M. K. Meade, J. O. Reagan, B. L. Byrnes, and M. Koohmaraie. 2001. Consumer impressions of Tender Select beef. *Journal of Animal Science* 79(10): 2605-2614.
- Signe, K., P. Freddy, and S. S. Erik. 1989. Method and apparatus for the determination of quality properties of individual cattle carcasses. European Patent No. 0321981.
- Smith, G. C. 1999. New technologies for precision selection, management and marketing of beef. In *Beef Information Days*. Pullman, WA: Washington State University.

- Steel, R. G. D., J. H. Torrie, and D. A. Dickey. 1997. *Principles and Procedures of Statistics - A Biometric Approach*. 3rd ed. New York, NY.: The McGraw-Hill Companies, Inc.
- Steiner, R., A. M. Wyle, D. J. Vote, K. E. Belk, J. A. Scanga, J. W. Wise, J. D. Tatum, and G. C. Smith. 2003a. Real-time augmentation of USDA yield grade application to beef carcasses using video image analysis. *Journal of Animal Science* 81(9): 2239-2246.
- Steiner, R., D. J. Vote, K. E. Belk, J. A. Scanga, J. W. Wise, J. D. Tatum, and G. C. Smith. 2003b. Accuracy and repeatability of beef carcass longissimus muscle area measurements. *Journal of Animal Science* 81(8): 1980-1988.
- Strang, G., and T. Nguyen. 1996. *Wavelets and Filter Banks*. Wellesley, MA: Wellesley-Cambridge Press.
- Stroshine, R. 1998. *Physical Properties of Agricultural Materials and Food Products*. IN.: West Lafayette: Copy Cat.
- Swatland, H. J., J. C. Brooks, and M. F. Miller. 1998. Possibilities for predicting taste and tenderness of broiled beef steaks using an optical-electromechanical probe. *Meat Science* 50(1): 1-12.
- Tong, A. K., D. J. Robinson, and T. Liu. 1998. Method and apparatus for using image analysis to determine meat and carcass characteristics. Canadian Patent No. 02263763.
- Unser, M. 1986. Sum and difference histograms for texture classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 8(1): 118-125.

- USDA. 1997. United States standards for grades of carcass beef. Effective date January 31, 1997. Washington, DC.: Livestock and Seed Division of the Agricultural Marketing Service. Available at:
<http://www.ams.usda.gov/lsg/stand/standards/beef-car.pdf>. Accessed on 9 February 2004.
- USDA. 2001. Procedures for approval and use of instrument augmentation systems for beef carcass ribeye measurement. Washington, DC.: Livestock and Seed Program, Agricultural Marketing Service.
- Van De Wouwer, G., P. Scheunders, and D. Van Dyck. 1999. Statistical texture characterization from discrete wavelet representations. *IEEE Transactions on Image Processing* 8(4): 592-598.
- Venel, C., A. M. Mullen, G. Downey, and D. J. Troy. 2001. Prediction of tenderness and other quality attributes of beef by near infrared reflectance spectroscopy between 750 and 1100 nm; further studies. *Journal of Near Infrared Spectroscopy* 9: 185-198.
- Vote, D. J., K. E. Belk, J. D. Tatum, J. A. Scanga, and G. C. Smith. 2003. Online prediction of beef tenderness using a computer vision system equipped with a BeefCam module. *Journal of Animal Science* 81(2): 457-465.
- Walker, J. 1996. Colour rendering of spectra. Switzerland: Fourmilab. Available at:
<http://www.fourmilab.ch/documents/specrend/>. Accessed on 15 February 2004.
- 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.

- Wheeler, T. L., and M. Koohmaraie. 1994. Prerigor and postrigor changes in tenderness of ovine longissimus muscle. *Journal of Animal Science* 72(5): 1232-1238.
- Wheeler, T. L., D. Vote, J. M. Leheska, S. D. Shackelford, K. E. Belk, D. M. Wulf, B. L. Gwartney, and M. Koohmaraie. 2002. The efficacy of three objective systems for identifying beef cuts that can be guaranteed tender. *Journal of Animal Science* 80(12): 3315-3327.
- Wheeler, T. L., L. V. Cundiff, and R. M. Koch. 1994. Effect of marbling degree on beef palatability in *Bos taurus* and *Bos indicus* cattle. *Journal of Animal Science* 72(12): 3145-3151.
- Wheeler, T. L., S. D. Shackelford, and M. Koohmaraie. 1998. Cooking and palatability traits of beef longissimus steaks cooked with a belt grill or an open hearth electric broiler. *Journal of Animal Science* 76(11): 2805-2810.
- Wheeler, T. L., S. D. Shackelford, and M. Koohmaraie. 2001. Shear force procedures for meat tenderness measurement. Clay Center, NE.: Roman L. Hruska U. S. Meat Animal Research Center, Agricultural Research Service, United States Department of Agriculture. Available at:
http://meats.marc.usda.gov/MRU_WWW/Protocol/WBS.pdf. Accessed on 17 February 2004.
- Winger, R. C., and C. J. Hagyard. 1994. Chapter 4: Juiciness - its importance and some contributing factors. In *Quality attributes and their measurement in meat, poultry and fish products. Advances in meat research series*. Vol. 9. 94-124. A. M. Pearson and T. R. Dutson, eds. London, UK.: Blackie Academic & Professional.

- Wulf, D. M., and J. K. Page. 2000. Using measurements of muscle color, pH, and electrical impedance to augment the current USDA beef quality grading standards and improve the accuracy and precision of sorting carcasses into palatability groups. *Journal of Animal Science* 78(10): 2595-2607.
- Wulf, D. M., and J. W. Wise. 1999. Measuring muscle color on beef carcasses using the L*a*b* color space. *Journal of Animal Science* 77(9): 2418-2427.
- Wulf, D. M., S. F. O'connor, J. D. Tatum, and G. C. Smith. 1997. Using objective measures of muscle color to predict beef longissimus tenderness. *Journal of Animal Science* 75(3): 684-692.
- Wyle, A. M., D. J. Vote, D. L. Roeber, R. C. Cannell, K. E. Belk, J. A. Scanga, M. Goldberg, J. D. Tatum, and G. C. Smith. 2003. Effectiveness of the SmartMV prototype BeefCam System to sort beef carcasses into expected palatability groups. *Journal of Animal Science* 81(2): 441-448.
- Wyle, A. M., R. C. Cannell, K. E. Belk, M. Goldberg, R. Riffle, and G. C. Smith. 1999. An evaluation of the prototype portable Hunterlab video imaging system (BeefCam) as a tool to predict tenderness of beef carcasses using objective measures of lean and fat color. 1999 Research Report. Fort Collins, CO.: Department of Animal Science, Colorado State University.
- Wyszecki, G., and W. S. Stiles. 1982. *Color Science. Concepts and Methods, Quantitative Data and Formulae*. 2nd ed. New York, NY.: John Wiley & Sons.
- Ye, J. 1993. A knowledge-based image analysis approach and its application in automatic beef grading. MS thesis. Guelph, Ontario: University of Guelph, Department of Computer Science.

APPENDICES

Appendix - A -- Development and Calibration of a Soft X-Ray Digital Imaging System for Agricultural Products.

APPENDIX A

DEVELOPMENT AND CALIBRATION OF A SOFT X-RAY DIGITAL IMAGING SYSTEM FOR AGRICULTURAL PRODUCTS

Nachiket Kotwaliwale, Jeyamkondan Subbiah,

Paul R. Weckler, Gerald H. Brusewitz, and Glenn A. Kranzler

Abstract

Soft X-ray imaging of agricultural products for quality determination is gaining worldwide interest. Recent advances in technology have made it possible to capture, store, and instantaneously process digital X-ray images at higher resolution for lower cost. An X-ray imaging system was developed and calibrated to capture images with a resolution of 1024 x 1024 pixels over a 50 mm x 50 mm area. Maximum X-ray tube voltage was 50 kVp, maximum current 1 mA, and signal integration time ranged from 460 to 6700 ms. The system consists of an X-ray source tube, a camera composed of a CMOS photodiode array, a frame grabber, and a data acquisition and control card. Because the X-ray beam is polychromatic, and the detector responses are variable, calibration is necessary to relate image intensity to X-ray attenuation. Imprecision of the system was 0.64% of the response range (1.63 gray level for an 8-bit pixel depth). Response, measured as mean pixel intensity, varied linearly with X-ray tube current and integration time and had a quadratic relationship with peak tube voltage. A regression

model was developed to estimate blank image intensity at higher voltages and currents. Prediction and validation errors for the model were 0.46% and 1.06% (1.18 and 2.72 gray level), respectively. Beam hardening effect was demonstrated using a polystyrene target. The procedure explained in this paper can be used to calibrate a soft X-ray imaging system for radiometric measurements.

Keywords. X-ray, Calibration, Digital radiograph, X-ray imaging, Beam hardening

Introduction

Non-destructive quality evaluation of agricultural products is a major area of interest for the agricultural processing industry. Availability of advanced technology has expanded avenues for non-destructive food quality determination. Techniques such as NIR imaging, MRI, and X-ray imaging have been explored to determine quality based on indicators not visible on the surface of the product (Chen and Sun, 1991). X-ray imaging is a leading contender to detect internal defects (Tollner, 2002). Successful applications include defect detection or quality evaluation of particle board (Steiner et al., 1978), pistachio nuts (Keagy et al., 1995), mangoes (Thomas, et al., 1995), apples (Schatzki et al., 1997), peaches (Barcelon et al., 1999), poultry (Tao and Ibarra, 2000), almonds (Kim and Schatzki, 2001), wheat (Karunakaran et al., 2002), and onions (Tollner, 2002).

Historically, X-ray imaging has been conducted on photographic plates or films (Curry et al., 1990), and internal defects were identified manually. Earlier applications of digital image processing algorithms on X-ray images required scanning of X-ray films (Keagy and Schatzki, 1993). Recent advances in technology have led to use of line-scan (Kim and Schatzki, 2000) and area-scan (Haff and Slaughter, 2002) cameras for X-ray

imaging. Improved digital hardware has enabled real-time X-ray imaging while reducing equipment cost.

Principle of X-ray Imaging

When high voltage in the order of kilovolts to megavolts is applied between electrodes in a vacuum, high-speed electrons from the cathode hit the anode of an X-ray tube, and X-rays are generated by Bremstrahlung and characteristic radiation processes (Curry et al., 1990). X-ray intensity depends on the quantity and quality of the radiation. Quantity refers to the number of photons, and quality refers to the energy of the photons. X-ray beam energy is the product of the number of photons and their associated energies. The quantity/flux of X-rays can be increased by increasing the X-ray tube current. Increasing the voltage between the two electrodes increases the mean photon energy in generated X-rays. Tube voltage limits the maximum energy of the photons. An X-ray beam produced by an X-ray tube has photons of different energies because of variation in energies of electrons hitting the anode and due to quantum effect i.e. the fact that most electrons give up their energy in stages (Curry et al., 1990). This polychromatic X-ray beam consists of a range of energies (wavelengths) of X-ray photons produced by the X-ray tube (Fig. A. 1). The spectrum in Figure 1 was simulated with X-ray toolbox (Siemens, 2001).

X-ray photons interact with orbital electrons of atoms of the material through which they are passing. In the soft X-ray region (<50 keV), X-rays photons are absorbed (“photoelectric collision”), transmitted, or scattered (Desrosier, 1960). Absorption and transmission are the most desired form of interaction in terms of achieving a quality image. Scattering creates an undesirable “fog” in the image. Fortunately, scattering

occurs for less than 10-15% of total interactions in the 10-50 keV range and therefore can be safely ignored (Curry et al., 1990).

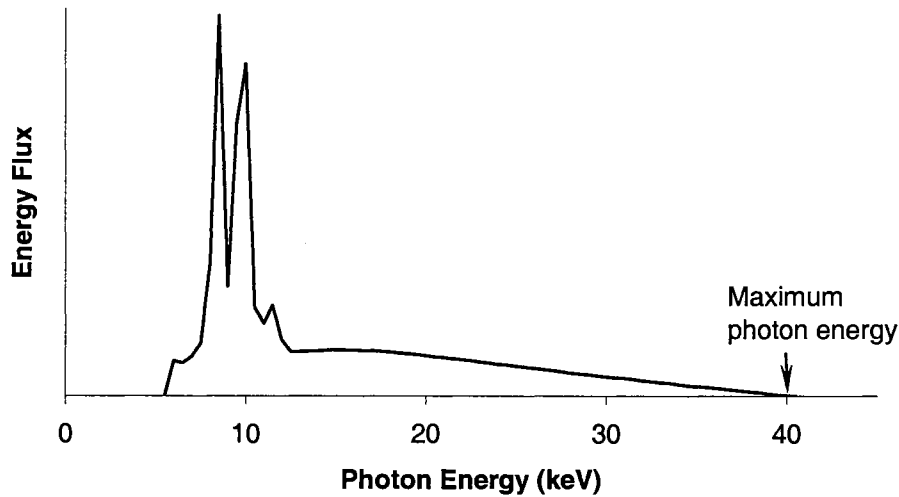


Figure A.1. A typical spectrum of X-rays produced by tungsten anode at 40 kV peak voltage with 125 μm Be filter (Siemens, 2001).

The total energy of an X-ray beam decreases exponentially as it moves through an object, due to the interaction of X-rays with matter. This phenomenon is called attenuation (Curry et al., 1990). For monochromatic radiation, the attenuation of X-rays through the material is given by:

$$I = I_0 \cdot \exp(-\mu_m \cdot t \cdot \rho) \quad (\text{A.1})$$

Where μ_m is mass attenuation coefficient (cm^2/g), ρ is material density (g/cm^3) and t is thickness (cm) through which the X-ray passes.

The mass attenuation coefficient is a property of the material and depends on atomic number and incident photon energy. If the absorbing material consists of more than one element, the mass attenuation coefficient of the composite material will be a function of mass attenuation coefficients of individual elements and their mass fraction in

the path of the photon beam. The difference between mass attenuation coefficient, density and thickness of two or materials in a composite forms image contrast in an X-ray image.

It is important to realize that this attenuation coefficient is a function of photon energy (Hubbell and Seltzer, 1996) and for monochromatic radiation it is constant for a material. For polychromatic X-rays each energy has its own coefficient value, thus the net attenuation coefficient of material changes with thickness of material. This phenomenon is known as “beam hardening” (Paiva et al., 1998). This effect causes an increase in the mean energy of the X-ray beam as it passes through the material.

Solid-state X-ray area detectors have spatial non-uniformities and distortions in the detector response. As a result, the recorded signal is not proportional to the incident X-ray intensity (Barna et al., 1999). These factors cause variability in detector-to-detector response, and therefore it is necessary that a digital X-ray imaging device be calibrated to relate detector response to the transmitted intensity of the X-ray beam.

An X-ray imaging system was developed to study X-ray attenuation of agricultural products and relate it to product quality. Calibration of the equipment is discussed in this paper.

Materials and Methods

Equipment

The equipment consists of an X-ray tube, X-ray camera, computer, digital frame grabber, and data acquisition and control card (Fig. A. 2).

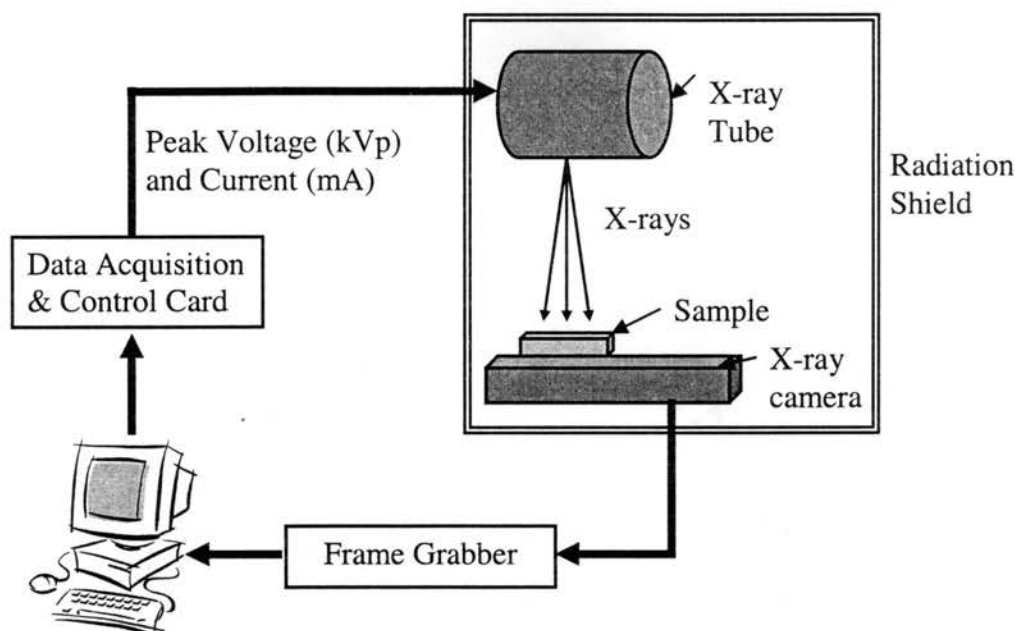


Figure A.2. Schematic of the equipment setup.

X-ray Tube

Suitable X-ray images of agricultural products can be acquired at X-ray tube voltages less than 50 keV (Steiner et al., 1978, Kim and Schatzki, 2000, Karunakaran et al., 2002). An X-ray tube (Model: XTF TM -5011, Oxford Instruments, X-Ray Technologies, Inc., Scotts Valley, CA), capable of operating at voltages from 4 to 50 kVp with maximum current of 1 mA was acquired. X-rays generated at a tungsten anode pass through 127- μm thick beryllium window with a diverging cone angle of 25° . The X-ray spot at the window exit is an oval of 76 x 93 μm .

X-ray Camera

An X-ray camera (Model: Shad-o-BoxTM -1024, Rad-Icon, Inc., Santa Clara, CA) was positioned below the target sample. The camera was constructed with a two-dimensional photodiode array of 1024 by 1024 pixels on 48- μm center-center spacing

giving a detector area of 49.2 x 49.2 mm. A Gd₂O₂S scintillator screen, placed in direct contact with the photodiode array, converts incident X-ray photons to visible light, which in turn are detected by the photodiodes. A graphite window blocks ambient light and protects the sensitive electronics from radiation damage. The analog signal from the photodiode array is digitized to 12-bit quantization in the camera and acquired by the frame grabber (Model: Imagenation[®] PXD 1000, Imagenation Corp., Beaverton, OR). The camera acquires image frames with signal integration time ranging from 460 to 6700 ms.

A data acquisition and control card (Model: Omega[®] DAQ 801 OM, Omega Engineering, Inc., Stamford, CT) was mounted in the computer ISA-bus for X-ray tube control. Software controlled peak voltage and current to the X-ray tube were implemented.

System Calibration and Testing

Calibration addressed the effect on detector response (X-ray image intensity) of: (a) the instrument characteristics, viz., dark current (camera) and polychromatic X-rays (source tube); and (b) the input variables, viz., signal integration time, X-ray tube current, and X-ray tube voltage. Images were captured with the camera placed 153 mm from the X-ray source and quantized at 8-bit depth for analysis. Only the central 1000 x 1000 pixels were used in this study because some pixels near the border were either dead or did not receive X-ray photons as expected. Images were analyzed using MATLAB (The Mathworks, Inc., Natick, MA), and all statistical analyses were conducted in SAS (SAS Institute, Inc., Cary, N.C.).

Dark Current

The detector array accumulates some signal even when not exposed to X-ray radiation. This effect is due to dark current (leakage) in the photodetector. Thermally generated charge is the main source of dark current in the photodetector array (Barna et al., 1999). Images generated by the photodetector consist of the signal due to dark current in addition to the signal generated by incident X-ray energy. Therefore, the dark current image (or offset image) must be deducted from all the images to obtain the net signal due to X-rays. To eliminate the effect of ambient temperature on signal or detector response, an offset image was acquired just before acquisition of each image used in the calibration study.

Flat-Field Correction

Non-uniformities in pixel response, phosphor scintillator, optical coupling, and transmission cause variation in response across the face of the detector (Gruner et al., 2002). Flat-field correction, also known as gain correction, is commonly employed to overcome these abnormalities and is multiplicative in nature (Barna, et al., 1999). The primary research interest was in calculating the attenuation coefficient, which is calculated from the ratio of two intensities, I and I_0 (Eqn. A.1). This multiplicative correction cancels out during division, and therefore flat-field correction was not employed on images used in this study.

Effect of Equipment Variables

To determine the effect of X-ray tube current on image detector response, images were acquired at currents from 0.1 to 1 mA in steps of 0.1 mA for voltages from 12.5 to 22.5 kVp in steps of 2.5 kVp, and from 25 to 50 kVp in steps of 5 kVp. Detectors

saturated at high voltages and currents (Ref. Section ahead – Development of Model) therefore, a suitable uniform material was placed over the detector array for voltages above 25 kVp such that the detectors did not saturate at 1 mA for that particular voltage. The lowest available integration time of 460 ms was used.

To study the effect of X-ray tube voltage on image detector response, blank images were acquired at currents from 0.2 to 1 mA in steps of 0.1 mA for voltages from 12.5 to 22.5 kVp in steps of 2.5 kVp. From 25 to 50 kVp in steps of 5 kVp, images were acquired at current levels of 0.25, 0.50, 0.75, and 1.00 mA with a suitable uniform material placed over detector so that the detectors did not saturate at 50 kVp for that particular current. Signal integration time was set to 460 ms.

To test the effect of signal integration time, images were acquired at 460, 640, 820, and 1000 ms for eight different combinations of voltage and current. A uniform material, if required, was placed over the detector, to prevent saturation at 1000 ms.

Repeatability of System

Imprecision is defined as lack of repeatability in the system. It consists of: (a) variability of the X-ray source tube in producing the desired quantity and quality of X-ray photons, and (b) variability in response of photodiode arrays to X-rays. An experiment was set-up as a factorial design with four levels of integration time (460, 640, 820, and 1000 ms), two levels of voltage supplied to tube (30, 50, kVp), and three levels of current supplied to tube (0.3, 0.6, and 0.9 mA). At each condition, if required, the X-ray beam was filtered using a uniform material to avoid image saturation. Three replicates were collected. Standard deviation of gray level at each pixel for three replicates was

calculated. Mean of standard deviation of gray level of all one million detectors (pixels) was reported as the imprecision of the system at that condition.

Development of Model

Blank images were acquired when no material was between the detector and X-ray tube. Blank images taken at different voltages and currents represent incident X-ray intensity (I_0) at that condition. I_0 is required to calculate attenuation coefficient (Eqn. A.1). Detector response increases with increasing voltage, current, or integration time, but reaches a limit after which an increase in any of these factors does not increase detector response. Figure A.3 shows a linear response for current and curvilinear response for voltage at lower voltage and current levels at 460-ms integration time. A plateau of surface at higher voltages and currents indicates saturation of detectors (Fig. A.3). Detector saturation complicates calculation of attenuation coefficient. For instance, at 50 kVp and 1 mA we should capture a good quality image of an agricultural product. Such an image would represent transmitted X-ray intensity (I) at 50 kVp and 1 mA. However, because the detector saturated at that condition, the estimate of I_0 is erroneous, and hence calculations of the attenuation coefficient would be incorrect. Therefore, calibration equations were developed to explain the relationship between voltage, current, and blank image intensity at 460 ms integration time. Detectors did not saturate at 1 mA and 460-ms for voltages below 25 kVp. Therefore, images were acquired at voltages from 25 to 50 kVp at intervals of 1 kVp, and at 10-15 different current levels, where detectors did not saturate. In total, 897 images were acquired to develop the model. The developed statistical model can be used to estimate extrapolated image intensity at higher voltages and currents, where detectors normally saturate.

Because detector-to-detector variation was quite high, (Fig. A.4), calibration equations were developed for each of the 1 million pixels.

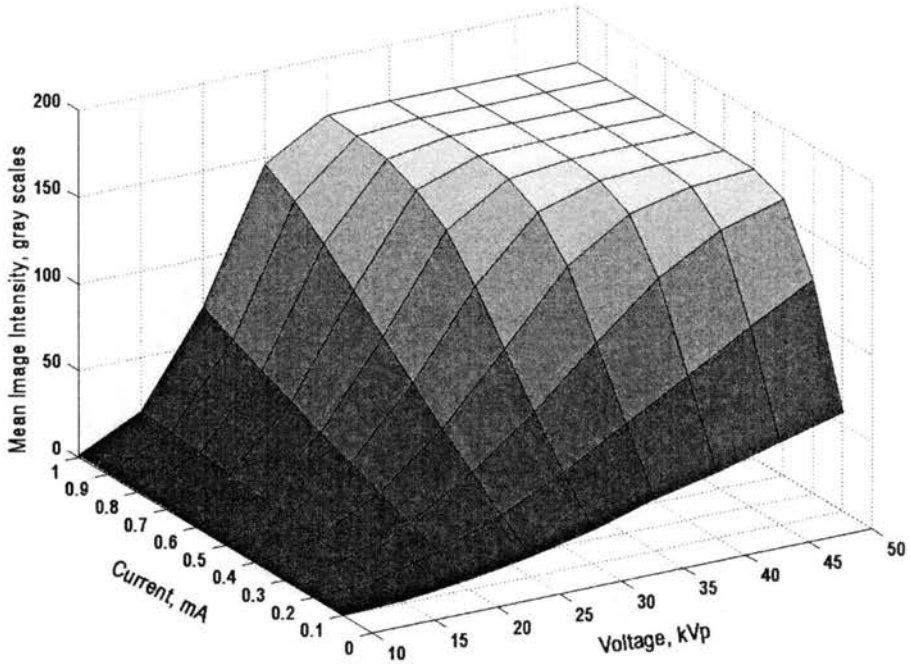


Figure A.3. Detector saturation at higher voltages and currents at 460 ms.

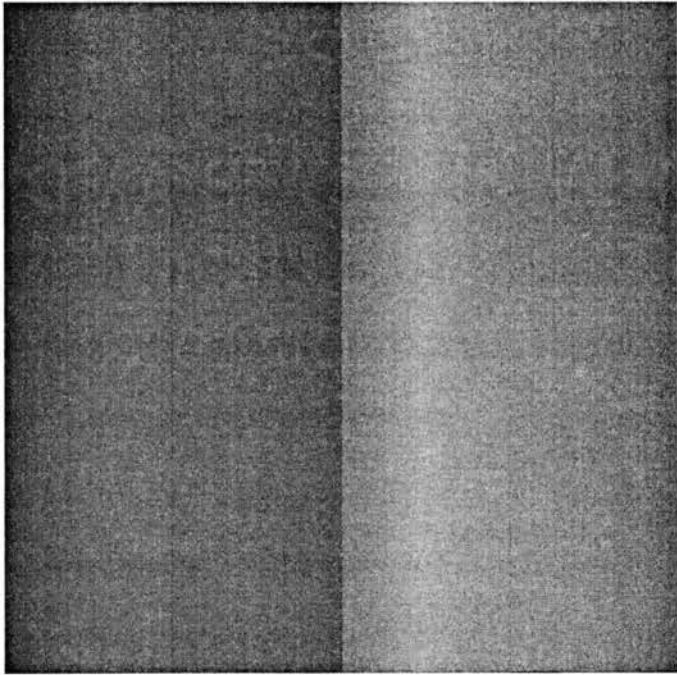


Figure A.4. Blank image (I_0) taken at 50 kVp, 0.22 mA, 460 ms demonstrating detector variation.

To validate the models, new blank images were acquired under conditions different from those with which the model was developed. Conditions were selected to avoid detector saturation. At each voltage from 25-50 kVp in steps of 1 kVp, one validation image was acquired.

Target Mass Attenuation Coefficient

Polystyrene sheets of thickness from 0.08 to 6.03 cm (22 levels) were used to predict mass attenuation coefficient (cm^2/g). Polystyrene was selected, because its density and atomic number are similar to biological materials (McFarlane, et al., 2000). Images were obtained at 460-ms signal integration time, at 15 to 50 kVp, and at three currents chosen such that the histogram distribution fell within the dynamic range of mean image intensities between 80 and 170 gray level. Blank images captured to develop statistical model were used for incident energy without attenuation (I_0), wherever possible. Otherwise, I_0 for the conditions were estimated from the developed model. Attenuation coefficient was calculated using Equation A.1.

Results and Discussion

Effect of X-ray Tube Current

Figure A.5 shows the scatter plot of current versus intensity at voltages from 12.5 kVp to 22.5 kVp. Intensity is linearly related to current with R^2 values of linear fit are above 0.99. Note that plots are for blank images, and thus the positive effect of voltage on intensity can also be seen. Scatter plots of current versus intensity at voltages from 25

to 50 kVp are shown in Figure A.6. However, voltage curves cannot be compared to each other, as different attenuation materials were used to avoid detector saturation.

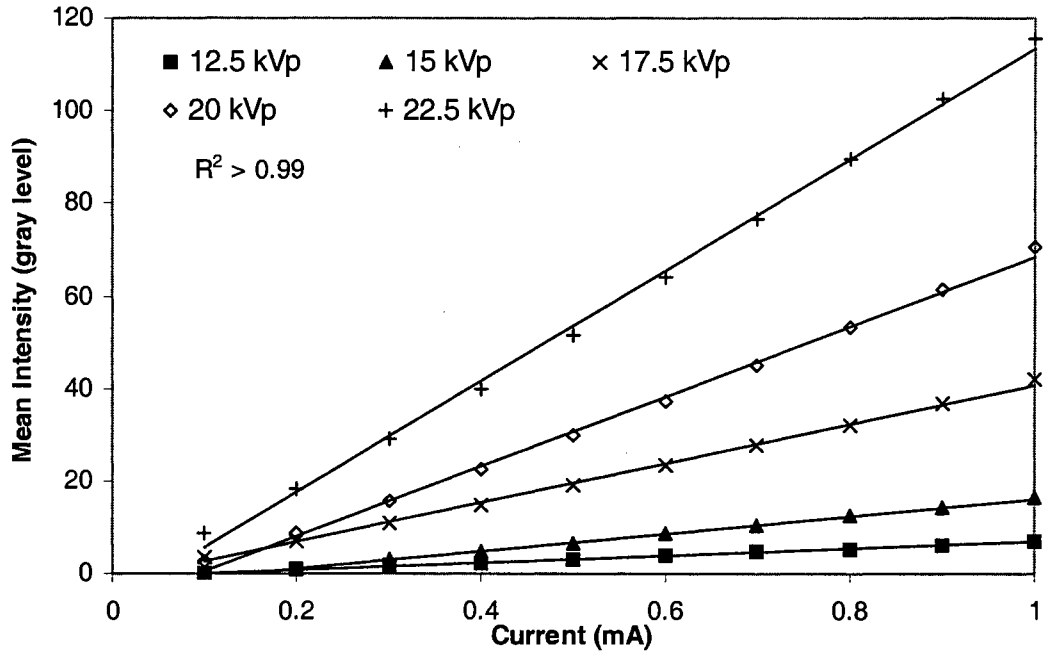


Figure A.5. Effect of current on intensity at lower voltages for blank images.

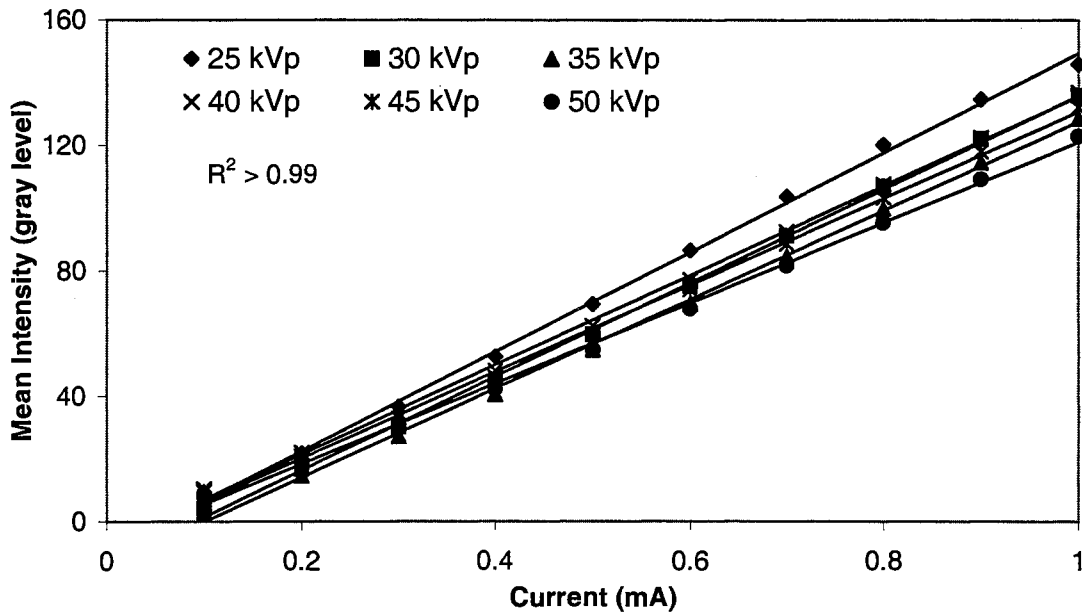


Figure A.6. Effect of current on intensity at higher voltages for images with filtered X-ray beam.

For instance, 0.127 mm mild steel sheet was used at 50 kVp whereas 3.175 mm LDPE and 4.76 mm HDPE sheet were used at 30 kVp. The attenuation materials were selected such that the intensity at the highest current (1 mA) was in the range of 120-160 gray level. Despite attenuation, at each voltage the linear effect of current on intensity is eminent. The linear relationship between current and image intensity is in agreement with the literature (Curry et al., 1990, Gambaccini et al., 1996).

Effect of X-ray Tube Voltage

The effect of voltage on intensity of blank images at voltages from 12.5 kVp to 22.5 kVp is presented in Figure A.7. Voltage showed a quadratic relationship with intensity. Peak voltage determines the maximum energy of the X-ray photons. In addition, higher voltage also increases the number of X-ray photons by increasing the electron cloud near the anode. Therefore, X-ray intensity increases as the square of the kilovoltage (Curry et. al., 1990). Regression models were constructed with intensity as dependent variable and voltage and square of voltage as independent variables. At lower currents (0.1 to 0.3 mA) the square of the voltage does not have a significant effect on intensity. At higher current (≥ 0.4 mA), the intensity depends on the square of the voltage. Adjusted R^2 values greater than 0.94 were obtained at various currents. Data at voltages higher than 22.5 kVp is not shown in Figure 7, because images at these voltages were taken with the X-ray beam filtered differently to avoid detector saturation and to keep the signal level high enough to detect. However, when analyzed for individual currents, at lower current (0.25 mA), the voltage-squared term was not significant. At higher currents (≥ 0.50 mA), the intensity was dependent on square of the voltage. Adjusted R^2 values greater than 0.99 were obtained at all currents. The quadratic effect

of voltage on intensity holds only for the unfiltered X-ray beam (Krestel, 1990). The beryllium window in the X-ray tube and filters used to avoid detector saturation can be identified as reasons for lack of a quadratic relation at lower currents. At higher currents the photon flux density was sufficient to show a significant quadratic relationship.

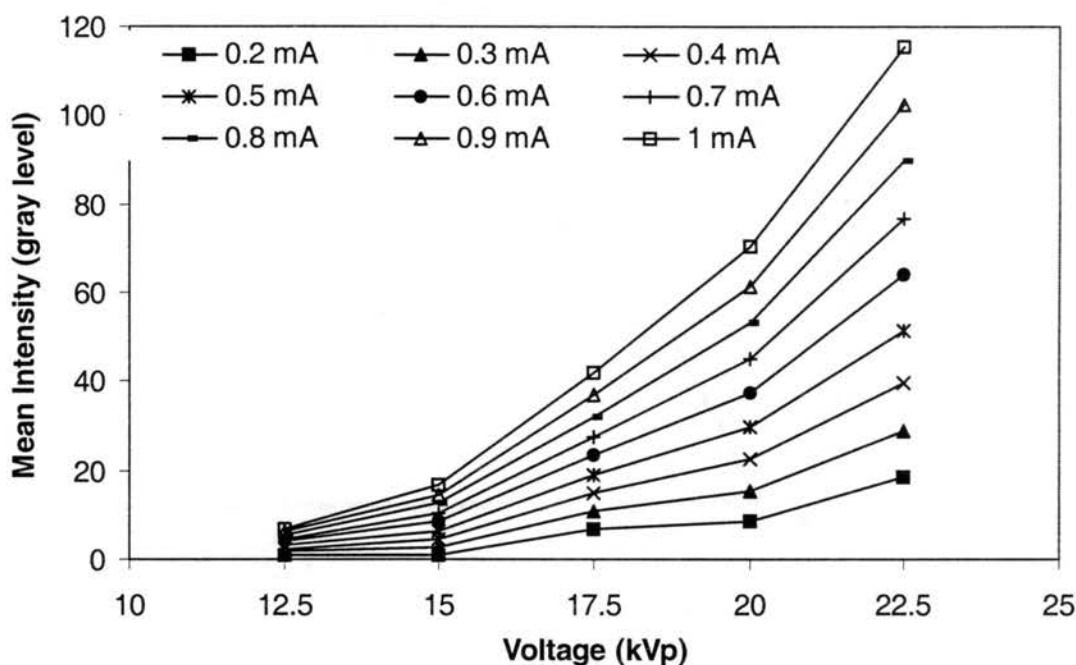


Figure A.7. Effect of voltage on intensity of blank images at various currents.

Effect of Integration Time

The effect of integration time on the mean image intensity is shown in Figure A.8. R^2 values of more than 0.99 were observed, indicating that the response varied linearly with signal integration time. Gambaccini et al., (1996) reported the cumulative effect of current and integration time to be linear on X-ray image intensity. Some deviations in linearity at high voltage levels can be attributed to detector saturation. Increasing integration time not only increases signal level but also increases dark current level. Detectors tend to saturate at longer integration times, causing a decrease in the net signal.

Figure A.9 shows this saturation effect for images taken at 15 kVp and 1 mA at different integration times.

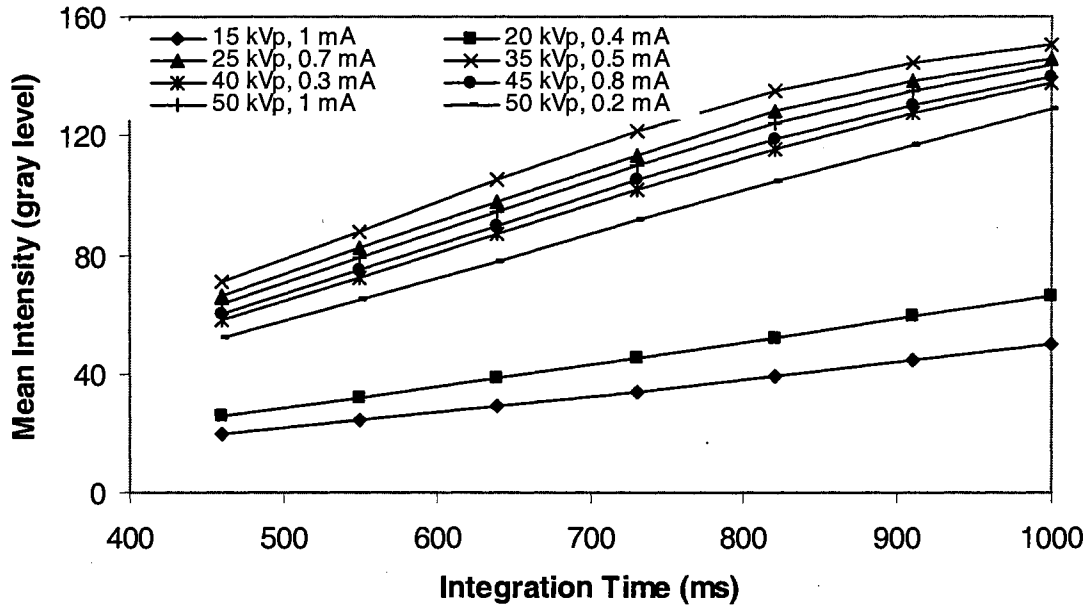


Figure A.8. Effect of integration time on the mean image intensity for eight different combinations of voltage and current.

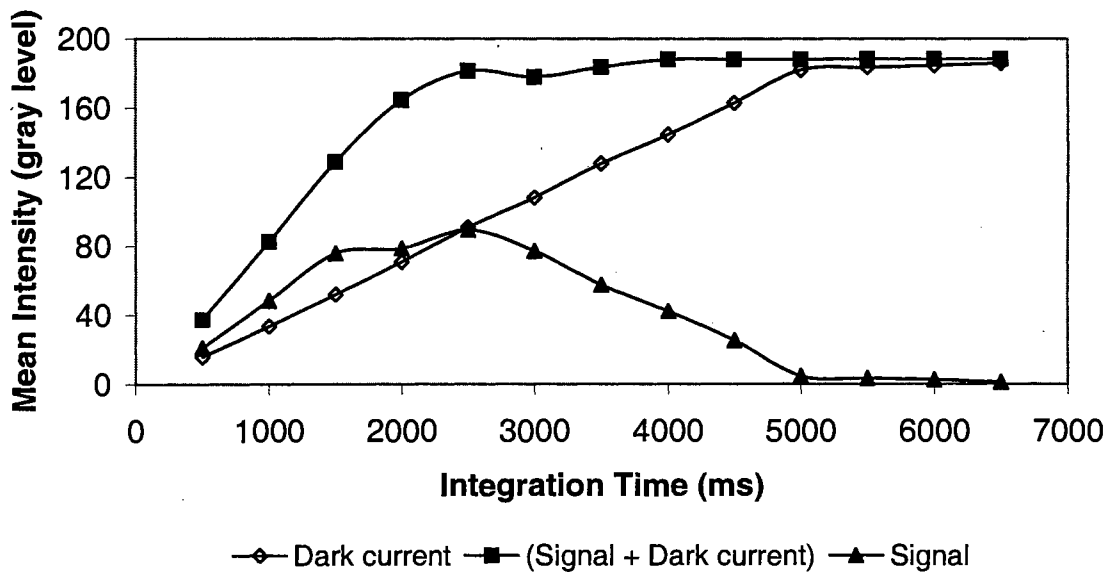


Figure A.9. Saturation in dark current and signal due to increase in integration time.

Integration times above 1500 ms result in loss of information and at 5000 ms both dark current and signal become saturated and hence the net signal drops to zero. The integration time limit for detector saturation would be different for different voltages.

Imprecision

System imprecision indicated by the mean standard deviation, ranged from 0.16% to 1.38% (0.41 to 3.52 gray level) with a mean of 0.64% (1.63 gray level). All combinations of two-way interaction of three factors were not significant at 95% level of significance. All factors, voltage ($p=0.15$), current ($p=0.22$), and integration time ($p=0.23$) did not have significant effect on the imprecision of the system. This result implies that the imprecision of the system is stable with respect to all input variables. The mean imprecision of 1.63 gray level is relatively insignificant compared to dynamic range (0-255 gray level).

Statistical Model

Statistical regression models were developed using data from blank images where detectors did not saturate. These models can be used to predict incident intensity at conditions in which detector saturation occurs. Because intensity is directly proportional to current (Fig. A.5 and A.6) and proportional to the square of the voltage (Fig. A.7), the following model was first evaluated.

$$I_0 = A + B \cdot i + C \cdot v^2 \quad (\text{A.6})$$

Where: I_0 is pixel intensity, A , B , and C are regression coefficients, i is current and v is tube voltage. When this model was applied, the adjusted R^2 value was only 0.68.

Because the intensity was linearly related to current, separate regression models were developed for each voltage level from 25-50 kVp in steps of 1 kVp. The statistical model evaluated was:

$$I_{0(v)} = A_v + B_v \cdot i \quad (A.7)$$

The R^2 values were greater than 0.99 for all models. The slope (B_v) in Equation A.7 was linearly related to voltage (Fig. A.10), as given by:

$$B_v = C + D \cdot v \quad (A.8)$$

By combining Equations A.7 and A.8, a single model was obtained to estimate pixel intensity (I_0) for all voltages (kVp) and currents (mA).

$$I_0 = A + C \cdot i + D \cdot i \cdot v \quad (A.9)$$

Coefficients A, C, and D were calculated for each of the one million pixels from 897 images captured for model development.

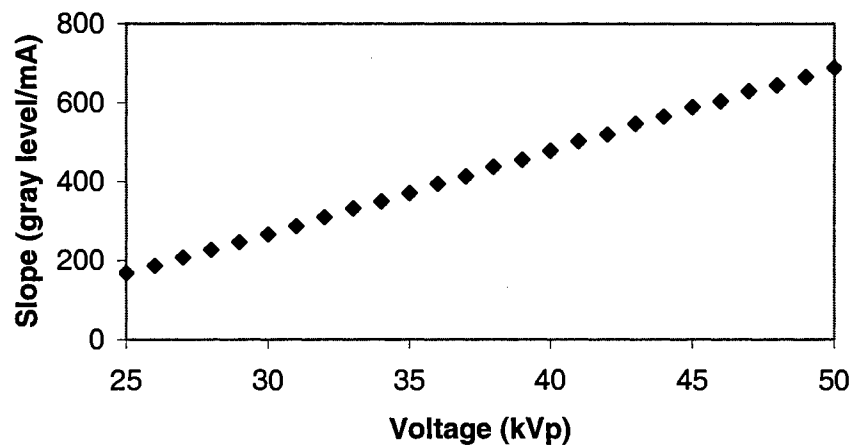


Figure A.10. Relationship between voltage and slope of the linear current model.

Evaluation of Statistical Models

The developed model (Eqn. A.9) was used to predict intensity of blank images at conditions under which the model was developed. Mean absolute deviation of the

predicted intensity from observed intensity values was 1.18 gray level (0.46% deviation). The predicted intensity versus observed intensity for different voltages and currents is shown in Figure A.11. Adjusted R^2 value was higher than 0.99 for a linear fit. The plot shows a slight sigmoid shape indicated that higher-order terms of (kVp*mA), mA, kVp could improve the degree of fit. Including higher-order coefficients did not improve the R^2 value nor the distribution residuals. The validation error was 1.06% (2.72 gray level). Errors are of the order of system imprecision (1.63 gray level), indicating the models are adequate for practical purposes.

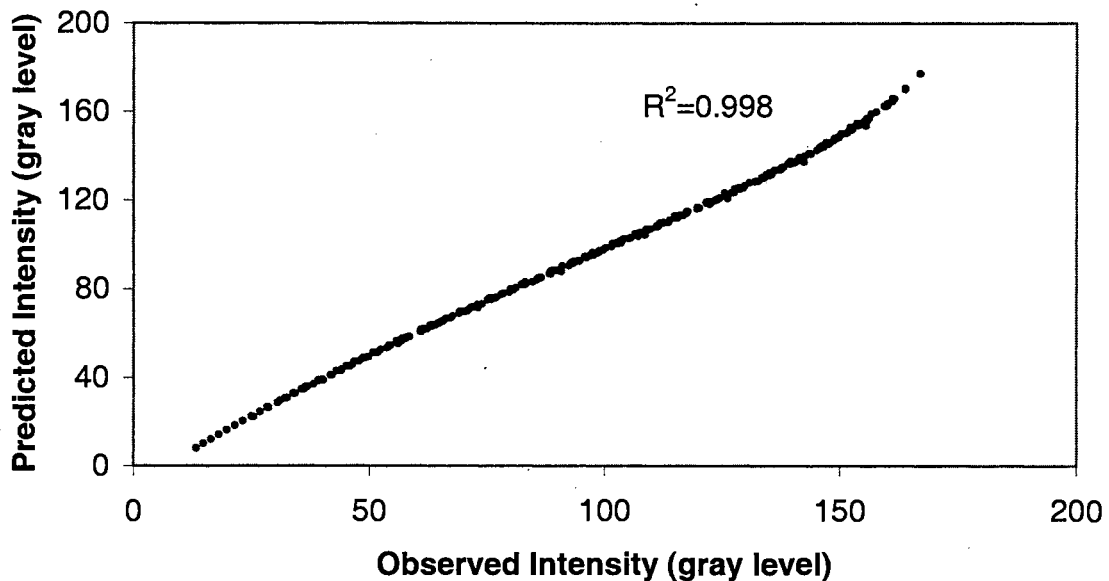


Figure A.11. Evaluation of model (Eqn. A.9) for calibration images.

Target Attenuation Coefficient

The measured mass attenuation coefficient of the polystyrene decreased with the thickness of polystyrene (Fig. A.12). This is due to “beam hardening” of polychromatic X-rays produced by the source tube. At a given thickness, mass attenuation coefficients were less for higher voltages (Fig. A.12) as also reported by Curry et al. (1990) and

Hubbell and Seltzer (1996). A sharp drop in attenuation coefficient values is evident at lower thickness. However, the drop decreased with thicker material. The attenuation coefficient values will drop until the polychromatic X-ray beam becomes effectively monochromatic, i.e. the mean energy of the beam reaches the applied peak voltage.

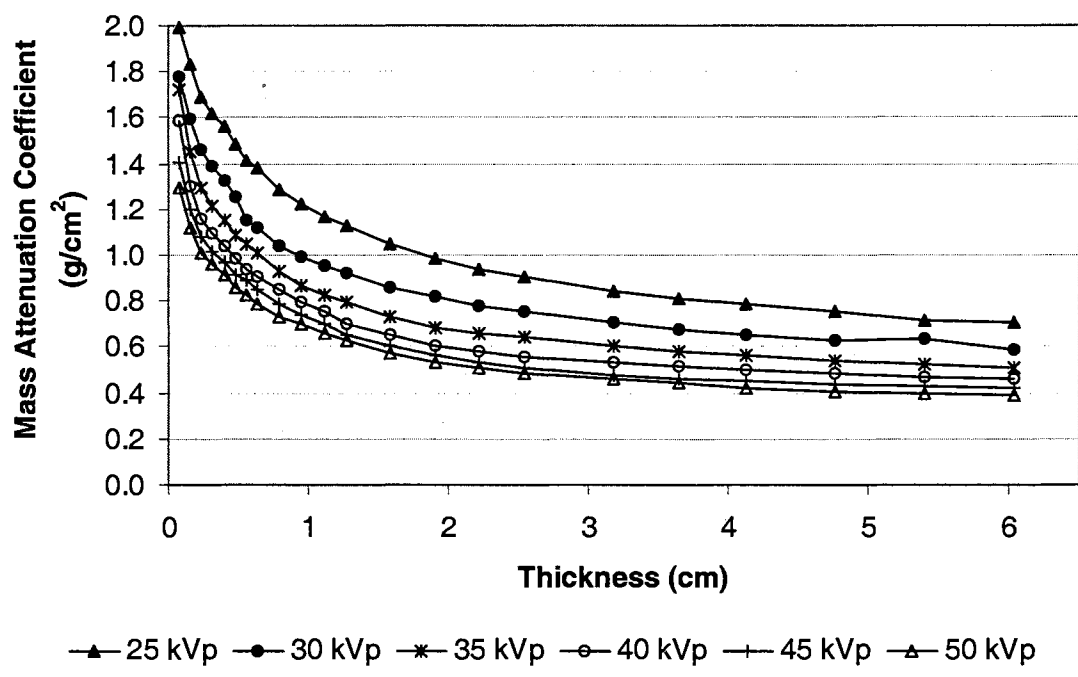


Figure A.12. Attenuation coefficient of polystyrene at different X-ray energies

Conclusions

Development of soft X-ray imaging system for agricultural products and calibration of that system was presented in this study. Repeatability of the system was established. Effect of X-ray imaging variables (X-ray tube voltage, current, integration time), inconsistencies in X-ray source and detectors and product thickness on the image pixel intensities was discussed in detail. An attempt was also made to deal with detector

saturation at higher voltages and currents. The following conclusions can be drawn from this study:

- X-ray tube current is linearly related to detector response, while X-ray voltage has a quadratic relationship.
- The signal integration time has a linear effect on detector response; however high integration times can saturate the detector.
- System imprecision ranged from 0.16% to 1.38% (0.41 to 3.52 gray level) with a mean of 0.64% (1.63 gray level).
- Statistical models developed to estimate incident intensity (I_0) at higher voltages and currents (at which sensors saturate) were successful with $R^2 > 0.99$, prediction error = 1.18 gray level, and validation error = 2.72 gray level. Error was of the order of system imprecision (1.62 gray level).
- Attenuation coefficient of polystyrene was determined, which varied with material thickness due to beam hardening effect.

Disclaimer

The use of trade names is for information purpose only and does not constitute endorsement by Oklahoma State University or the authors.

Acknowledgements

A USDA special grant and the Oklahoma Agricultural Experiment Station supported this research.

References

- Barcelon, E. G., S. Tojo, and K. Watanabe. 1999. X-ray Computed Tomography for Internal Quality Evaluation of Peaches. *Journal of Agricultural Engineering Research* 73(4): 323-330.
- Barna, S. L., M. W. Tate, S. M. Gruner, and E. F. Eikenberry. 1999. Calibration procedures for charge-coupled device x-ray detectors. *Review of Scientific Instruments* 70(7): 2927-2934.
- Chen, P., and Z. Sun. 1991. A review of non-destructive methods for quality evaluation and sorting of agricultural products. *Journal of Agricultural Engineering Research* 49: 85-98.
- Curry, T. S., J. E. Dowdey, and R. C. Murry. 1990. *Christensen's Physics of Diagnostic Radiology*. 4th ed. Media, PA.: William & Wilkins.
- Desrosier, N. W. 1960. *Radiation Technology in Food, Agriculture, and Biology*. Westport, CT.: AVI Publishing Company.
- Gambaccini, M., A. Taibi, A. Del Guerra, M. Marziani, and A. Tuffanelli. 1996. MTF evaluation of a phosphor-coated CCD for x-ray imaging. *Physics In Medicine And Biology* 41(12): 2799-2806.
- Gruner, S. M., M. W. Tate, and E. F. Eikenberry. 2002. Charge-coupled device area x-ray detectors. *Review of Scientific Instruments* 73(8): 2815-2842.
- Haff, R. P., and D. C. Slaughter. 2002. X-ray inspection of wheat for granary weevils. Real time digital imaging vs. film. ASAE Paper No. 02-6093. St. Joseph, Mich.: ASAE.

- Hubbell, J. H., and S. M. Seltzer. 1996. NISTIR 5632: Tables of X-Ray Mass Attenuation Coefficients and Mass Energy-Absorption Coefficients from 1 keV to 20 MeV for Elements $Z = 1$ to 92 and 48 Additional Substances of Dosimetric Interest. Gaithersburg, MD.: Ionizing Radiation Division, Physics Laboratory, National Institute of Standards and Technology. Available at: <http://www.physics.nist.gov/PhysRefData/XrayMassCoef/cover.html>. Accessed on 24 July 2003.
- Karunakaran, C., D. S. Jayas, and N. D. G. White. 2002. Soft X-ray Inspection of wheat kernels infested by *Sitophilus oryzae*. ASAE Paper No. 02-3132. St. Joseph, Mich.: ASAE.
- Keagy, P. M., and T. F. Schatzki. 1993. Machine recognition of weevil damage in wheat radiographs. *Cereal Chemistry* 70(6): 696-700.
- Keagy, P. M., B. Parvin, and T. F. Schatzki. 1995. Machine recognition of navel orange worm damage in x-ray images of pistachio nuts. In *Proceedings of SPIE*, 2345: 192-203. Bellingham WA.: The International Society for Optical Engineering.
- Kim, S., and T. F. Schatzki. 2000. Apple water-core sorting system using X-ray imagery: I. Algorithm development. *Transactions of the ASAE* 43(6): 1695-1702.
- Kim, S., and T. Schatzki. 2001. Detection of pinholes in almonds through x-ray imaging. *Transactions of the ASAE* 44(4): 997-1003.
- Krestel, E. 1990. *Imaging Systems for Medical Diagnosis: Fundamentals and Technical Solutions - X-Ray Diagnostics- Computed Tomography - Nuclear Medical Diagnostics - Magnetic Resonance Imaging - Ultrasound Technology*. Indianapolis, IN: John Wiley & Sons, Inc.

- McFarlane, N. J. B., C. R. Bull, R. D. Tillett, R. D. Speller, G. J. Royle, and K. R. A. Johnson. 2000. The Potential for Compton Scattered X-rays in Food Inspection: The Effect of Multiple Scatter and Sample Inhomogeneity. *Journal of Agricultural Engineering Research* 75(3): 265-274.
- Paiva, R. F. D., J. Lynch, E. Rosenberg, and M. Bisiaux. 1998. A beam hardening correction for X-ray microtomography. *NDT & E International* 31(1): 17-22.
- Schatzki, T. F., R. P. Haff, R. Young, I. Can, L. C. Le, and N. Toyofuku. 1997. Defect detection in apples by means of X-ray imaging. *Transactions of the ASAE* 40(5): 1407-1415.
- Siemens. 2001. Simulation of X-ray spectra - Radiography. Erlangen, Germany: Vacuum Technology Division, Siemens Medical Solutions. Available at: <http://www.med.siemens.com/med/rv/spektrum/mamIn.asp>. Accessed on 29 May 2003.
- Steiner, P. R., L. A. Jozsa, M. L. Parker, and S. Chow. 1978. Application of X-ray densitometry to determine density profile in waferboard: Relationship of density to thickness expansion and internal bond strength under various cycles. *Wood Science* 11(1): 48-55.
- Tao, Y., and J. G. Ibarra. 2000. Thickness-compensated X-ray imaging detection of bone fragments in deboned poultry - Model analysis. *Transactions of the ASAE* 43(2): 453-459.
- Thomas, P., A. Kannan, V. H. Degwekar, and M. S. Ramamurthy. 1995. Non-destructive detection of seed weevil-infested mango fruits by X-ray imaging. *Postharvest Biology and Technology* 5(1-2): 161-165.

Tollner, E. W. 2002. Classification of onions based on internal defects using commercial X-ray inspection equipment. ASAE Paper No. 02-6092. St. Joseph, Mich.: ASAE.



VITA

Jeyamkondan Subbiah

Candidate for the Degree of

Doctor of Philosophy

Dissertation: NONDESTRUCTIVE EVALUATION OF BEEF PALATABILITY

Major Field: Biosystems Engineering

Biographical:

Personal Data: Born in Madurai, India in 21 March 1975 to Saraswathy and Subbiah. Married to Latha in 6 December 2000. Son, Vishal born in 29 September 2003 in Stillwater, OK.

Education: Graduated from Thiyagarajar Model Higher Secondary School, Madurai, India in 1992; received Bachelor of Engineering in Agriculture from Tamil Nadu Agricultural University, Kumulur, India in 1997; received Master of Science in Biosystems Engineering from University of Manitoba, Winnipeg, Canada in 1999. Completed the requirements for the Doctor of Philosophy degree with a major in Biosystems Engineering at Oklahoma State University in May, 2004.

Experience: Worked as a Graduate Research Assistant at Department of Biosystems Engineering, University of Manitoba, Sept. 1997 – Dec. 1999; employed as a Research Engineer by the Department of Biosystems and Agricultural Engineering, Oklahoma State University, Jan. 2000 – present.

Membership in Honorary Societies: Sigma-Xi, The Scientific Honorary Society; Phi Kappa Phi, The National Honorary Society; Tau Beta Pi, National Engineering Honor Society; Alpha Epsilon, The Honor Society of Agricultural, Food, and Biological Engineering.

Membership in Scientific Societies: American Society of Agricultural Engineers; Institute of Food Technologists; Society of Optical Engineering.

Professional Registration: Engineer-in-Training of Oklahoma State Board of Professional Engineers and Land Surveyors, OK and Association of Professional Engineers & Geoscientists of the province of Manitoba, Canada.