Colour Image Processing and Texture Analysis on Images of Porterhouse Steak Meat

Lee Streeter[1,2], Robert Burling-Claridge[2], Michael J. Cree[1]

[1]University of Waikato, Dept. Physics and Electronic Engineering, Hamilton.

[2]AgResearch, Ruakura Research Centre, Hamilton

Email: lvs2@phys.waikato.ac.nz

Abstract

This paper outlines two colour image processing and texture analysis techniques applied to meat images and assessment of error due to the use of JPEG compression at image capture. JPEG error analysis was performed by capturing TIFF and JPEG images, then calculating the RMS difference and applying a calibration between block boundary features and subjective visual JPEG scores. Both scores indicated high JPEG quality. Correction of JPEG blocking error was trialled and found to produce minimal improvement in the RMS difference. The texture analysis methods used were singular value decomposition over pixel blocks and complex cell analysis. The block singular values were classified as meat or non-meat by Fisher linear discriminant analysis with the colour image processing result used as 'truth.' Using receiver operator characteristic (ROC) analysis, an area under the ROC curve of 0.996 was obtained, demonstrating good correspondence between the colour image processing and the singular values. The complex cell analysis indicated a 'texture angle' expected from human inspection.

Keywords: Meat Imaging, Colour Image Processing, Singular Value Texture Analysis, Complex Cell Analysis.

1 Background

In this paper we investigate the relationship of texture analysis segmentation to colour segmentation on meat images and investigate image features to characterise visual texture in an image.

The appearance of meat, such as colour and visual texture are important to the consumer when purchasing meat. The perceived colour of meat is a function of its chemical makeup and viewing conditions, varying with age and ambient lighting. Visual texture aspects include the alignment of muscle fibres with respect to the axis of view and "marbling" or the distribution of small fat regions in the muscle proper; the latter feature being particularly important in certain markets. Segmentation of meat and fat from the background, meat from the outer fat and meat from marbling by colour are conceptually easy tasks provided certain constraints, such as controlled lighting and positioning are achieved. Texture analysis on meat on the other hand is somewhat more subtle.

The Meat Quality project at AgResearch is a FRST funded initiative with the ongoing goal of finding better ways to objectively assess meat for features salient to the consumer such as colour, texture (to the bite), firmness and moisture. Visible/near infrared spectroscopy (VIS/NIR) is being employed as a possible commercial tool for nondestructively measuring the chemical and physical properties of meat. VIS/NIR spectroscopy has traditionally only scanned a single point or small region at a time. Spectroscopic imaging (usually referred to as hyperspectral imaging in the literature) provides spectral information at a number of localisations, yielding considerable extra information. Hyperspectral imaging systems are being developed, but as yet commercially available systems either have limited spectroscopic bandwidth or considerably limited spectral resolution (these cases are usually referred to as multispectral imaging). Regardless, given a suitable hyperspectral imaging device, there is considerable room for research into the fusion of traditional NIR spectroscopic data analysis and image processing techniques. Meanwhile it suffices to investigate processing on pseudo-hyperspectral images to consider the necessary data processing methods.

As part of the Meat Quality research, AgResearch is looking at localised properties throughout the volume of Longissimus Dorsi (LD) muscles (porterhouse or ribeye) by NIR, a non-invasive technique. This "3-D mapping" is hoped to reveal new information about spatial variability in the characteristics of the muscles. Image processing is being investigated to establish localised 'truth' of meat

content (fat-muscle ratios, etc) for the NIR spectral analysis. This paper describes some image processing investigated.

This paper is structured as follows. Section 2 outlines the methodology used, including image acquisition, steps taken to assess JPEG quality, colour image processing and texture analysis. Section 3 gives results and discussion and Section 4 conclusions and future directions.

2 Methodology

2.1 Setup and Image Acquisition

A Longissimus Dorsi (porterhouse or ribeye) muscle was cut from a beef carcass following a very careful handling regime, rolled in plastic wrap and stored at 15^0C post rigor for three days. The wrapped meat was then placed in a tube holder to ensure the surface was perpendicular to the intended view angle and to allow clean cutting as needed. A large blue board was placed behind the meat from the intended view of the camera to assist in image segmentation. Slices were spaced at 14mm separation along the axis of the muscle and were taken immediately prior to image acquisition to minimise colour effects due to the chemical interaction of oxygen.

Camera setup:

- A Sony DSC707 commercial digital camera was used.
- The camera was set up with the camera lens aligned 0.9 m from the meat face. The camera zoom was set to maximal optical zoom and the focus set at 0.9 m.
- The depth of focus was set to the minimum possible so that non-target background objects were blurred.
- Camera resolution was set to the maximum resolution 2560×1920 pixels.

Image Acquisition:

- Three images of each meat slice with one grey 18% reflectance and one white 90% reflectance image were taken (Jessops photographic grey card) as reference grey for each meat slice.
- Six slices were studied giving a total of 30 images.
- For the assessment of the quality of JPEGs produce by the camera, JPEG and TIFF copies of each image were generated by the camera. The TIFF to JPEG compression ratio was 6.5:1.

2.2 JPEG Quality Assessment

The JPEG file format was chosen because of the time taken to save an individual lossless TIFF file on the camera (approximately 45 seconds for each 14 Mbyte TIFF file) and due to the need to aquire data quickly over the slices of a muscle to ensure minimal time influence on the chemical composition of the muscle. To assess the error due to JPEG compression, the RMS difference for the individual colour planes was taken over the image set. Visual inspection of the JPEGs showed that significant blocking only occurred in the blurred blue background region of the images. No ringing artefacts were observed.

The JPEG deblocking algorithm due to Pham and van Vliet [1] was trialled. This algorithm uses anisotropic diffusive filtering by convolution with a variable linear Gaussian element to smooth along JPEG blocks while attempting to preserve image edges. The result combined with an isotropic global Gaussian filtered image by a weighted addition soft mixing method. Automatic JPEG quality assessment [2] was used as suggested in [1] to set the strength of filtering. The quality assessment scores JPEGs between 0 (worst visual quality) and 10 (best visual quality) with the worse the quality the greater the standard deviation of the Gaussian elements.

2.3 Colour Image Processing

Figure 1 shows an outline of the colour image processing algorithm used to segment the meat images into the constituent parts of the meat region, large fat regions and the blue background. A simplified L*a*b space was used [3]

$$\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

where c_1 is 'brightness', c_2 provides a red-enhanced image useful for detecting the meat region and c_3 a blue-enhanced image.

The blue region was detected by thresholding c_3 at the threshold value 0 and the threshold result binary image was morphologically opened using a circular structuring element of radius 3. The meat region was detected in c_2 by first morphologically closing with a square structuring element of width 5 pixels to reduce small dark regions in the meat area. Then the Haar wavelet transform was applied until a minimum of eight rows was reached. In the minimum scale image the maximum intensity pixel point was found and tracked back up through the subsequent resolutions in the approximation

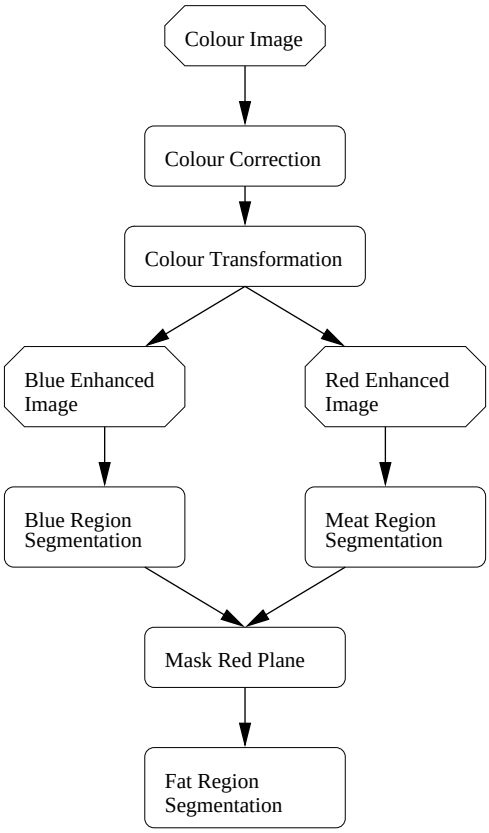

Figure 1: Block Diagram of Colour Processing

coefficients to the original resolution. The closed c_2 image was then thresholded at the mean plus the standard deviation of all pixel values in the image. Holes in the resultant binary image were flood-filled. The object determined by the corse to fine resolution tracking was selected as the meat object.

The fat region was found by taking the red plane, masking out the meat and blue regions and then thresholding with a stringent (to avoid confusion with other bright objects) threshold value slightly less that the maximum pixel intensity value. The resultant binary image was morphologically closed using a circular element and then only objects greater than 1000 pixels in area that neighbour the meat region by 5 pixels were kept as fat objects.

2.4 SVD Texture Analysis and Classification

The meat region of the images exhibit a visually rough and reasonably regular textural pattern. Singular valued decomposition was investigated as a likely method for detecting the meat region.

Singular value decomposition (SVD) of matrices can be a useful discriminating factor between textures [4]. The singular value decomposition of a matrix X is

$$X = USV^T$$

where U and V are orthogonal matrices and S is a diagonal matrix with the singular values s_{ii} on the diagonal. The singular values are ranked $s_{i,i} > s_{i+1,i+1}$ and decay with respect to i at a rate proportional to the degree of variation in a matrix.

Singular values were computed over blocks of size 32×32. The block size was chosen as a multiple of eight to minimise any effects due to JPEG block boundaries. A block size of greater than eight pixels square was required to encompass enough textural information to be useful for texture discrimination, yet small enough for sufficient localisation. Blocks were begun in the top lefthand corner and left over rows and columns were discarded as the meat region is centred in the images, resulting in $40 \times 60 \times 32$ SVD images. When applied to a colour image and the singular values catenated we obtained $40 \times 60 \times 96$ SVD images for comparison with the colour image processing.

Pattern classification was applied to the singular values using Fisher linear discriminant analysis (LDA) [5]. Fisher LDA attempts to find the optimal solution to the linear classification problem

$$y_i = \sum_j x_{i,j} b_j \tag{2}$$

where $y_i = \{0, 1\}, \forall i$ is the set of class assignments corresponding to each block i, $x_{i,j}$ is the data array with singular values j and b_j are the classification coefficients. The coefficients b_j form a discriminating hyperplane in dataspace to be optimised by maximising the signal to noise ratio (written in matrix form)

$$\mathrm{SNR}(\mathbf{b}) = \frac{\mathbf{b}^T S_b \mathbf{b}}{\mathbf{b}^T S_w \mathbf{b}} \tag{3}$$

with the between class scatter matrix

$$S_b = (\mu_1 - \mu_2)^T (\mu_1 - \mu_2)$$

and the within class scatter matrix

$$S_w = \sum_c (X_c - \mu_c)^T (X_c - \mu_c)$$

where $c = \{0, 1\}$ is the designation of the two pattern classes with data X_c and mean μ_c over instances in class c. Equation 3 is maximised when

$$\mathbf{b} = S_w^{-1} (\mu_1 - \mu_0) \tag{4}$$

The SVD instances were each assigned a classification as meat ($y_i = 1$) or non-meat ($y_i = 0$) designating a block a meat block if at least half of all pixels in the meat segmented image from the colour segmentation above were set to one.

Receiver operator characteristic (ROC) analysis [6] was employed to examine the efficacy of texture classification. A ROC graph is a plot of, for a number of classification thresholds, the true positive fraction (TPF) or ratio of the number of true positives to number of actual positives against the false positives fraction (FPF) or ratio of false positive to actual negatives. Assuming the data follows the bi-normal parametric model the area under the ROC curve (AUC) is related to the signal to noise ratio of classification by

$$\text{SNR} = 2\text{erf}^{-1}(2\text{AUC} - 1) \tag{5}$$

where $\text{erf}^{-1}(\cdot)$ is the inverse error function. When AUC = 0.5, SNR = 0 which corresponds to noise and AUC = 1 implies TPF = 1 when FPF = 0, corresponding to perfect classification or SNR = ∞.

2.5 Complex Cell Operator for Image Texture Orientation Assessment

Figure 2: Segment of meat image

Texture in the meat images is visibly orientated with a clear dominant angle. Figure 2 shows an example segment of a meat image with diagonally directed texture from top left to bottom right. This angle of texture varies down the length of a muscle due to the change in direction of muscle fibres, influencing various subjective features of meat quality such as 'bite texture.' NIR spectra taken down the length of muscle fibrils are not likely to be quite the same as spectra taken across fibrils due to changes in light scattering.

To characterise texture orientation and strength of orientation a complex cell model is used [7]. The complex cell model is a neurologically inspired bar and edge detector intended to respond primarily, but not exclusively, to pixel intensity change in a given direction. To compute the complex cell response, first the Gabor wavelet response was taken where the gabor filter is:

$$\begin{aligned} g_{\xi,\eta,\lambda,\theta,\phi}(x,y) &= exp\left(-\frac{x^2+\gamma y^2}{2\sigma^2}\right)\cos\left(2\pi\frac{x}{\lambda}+\phi\right) \\ x &= (x_0-\xi)\cos\theta - (y_0-\eta)\sin\theta \\ y &= (x_0-\xi)\sin\theta + (y_0-\eta)\cos\theta \end{aligned} \tag{6}$$

with $\gamma = 0.5$. The response of image I to g was

$$r_{\lambda,\theta,\phi}(\xi,\eta) = \iint_\Omega I(x,y)g_{\xi,\eta,\lambda,\theta,\phi}(x,y)\,dxdy \tag{7}$$

where Ω was the image domain. An approximation image $a_{\xi,\eta,\lambda}$ was also computed as

$$a_\lambda(\xi,\eta) = \iint_\Omega I(x,y)e^{-\frac{(x-\xi)^2+\gamma(y-\eta)^2}{2\sigma^2}}\,dxdy \tag{8}$$

From r and a the simple cell was computed as

$$s_{\xi,\eta,\lambda,\theta,\phi}(x,y) = \begin{cases} 0 & \text{if } a = 0 \\ H(\frac{Rr}{a}/(\frac{r}{a}+C)) & \text{otherwise} \end{cases} \tag{9}$$

where $H(\cdot)$ is the Heavyside function, R and C are free parameters that control the maximum response and the semi-saturation constant respectively. R and C were set to 1 for this study. Finally the complex cell response was found from the simple cell response as

$$c_{\xi',\eta',\lambda,\theta} = \iint_\Omega e^{-\frac{(x-\xi')^2+\gamma(y-\eta')^2}{2(2\sigma)^2}}\sqrt{\sum_\phi s^2_{\xi,\eta,\lambda,\theta,\phi}}\,d\xi d\eta \tag{10}$$

which is a diffused root-sum-squares of the simple cell responses with preferred orientation and spatial frequency.

With the complex cell response in hand, we may now attempt to find an indication of the strength of each preferred angle response over the meat region. This angular response strength measure was hoped to serve as an indication of muscle fibril orientation when considering NIR analysis. We mask out all non-meat regions of the image and average the complex cell responses over 32×32 blocks to match the SVD texture analysis. Then for each spatial frequency at each block the angle for which the block has the greatest response was found, indicating the dominant textural direction for each block. For each angle the number of dominant blocks under the meat region was plotted as an average complex cell response verses angle graph. Visually the green plane showed the greatest intensity variation in the meat region and thus was chosen for complex cell response computation.

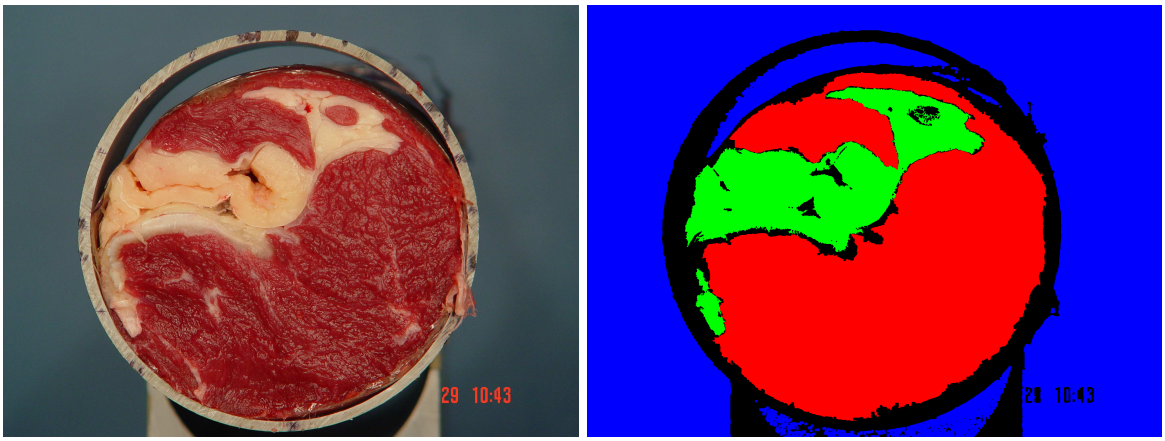

Figure 3: (left) A meat image and (right) the colour segmented image

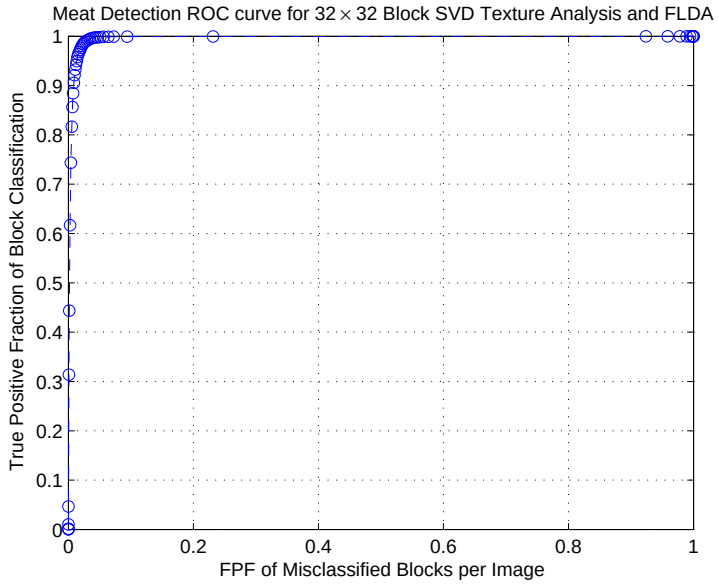

Figure 4: ROC curve for classifying blocks as meat by singular values using Fisher LDA. AUC = 0.996

3 Results and Discussion

The mean RMS JPEG error over the thirty images (with intensity range $(0, 255)$) was 1.83 (standard deviation 0.42), after deblocking the mean error was 1.71 (standard deviation 0.80). Thus the original RMS error was only 0.72 of pixel intensity. However a paired t-test at the 95% confidence limit rejected the hypothesis that the mean RMS improvement was statistically insignificant (however small). Concordant with the low RMS intensity error is visual assessment of the JPEGs shows ringing artefacts appear visually non-existent and that the JPEG visual quality assessment algorithm [2] gave a ten out of ten score for all images.

Figure 3 shows a meat image and the result of the colour segmentation into meat, fat and blue background regions. Unclassified regions in the meat face are mostly connective tissue or small 'disjoint' meat regions. This particular slice is an example result of incorrect slicing technique which (according to an AgResearch expert butcher) is due to "pressing down on the meat like a loaf of bread while cutting."

Figure 4 shows the ROC curve for classifying image blocks as meat by SVD texture analysis using the colour image processing result as 'truth.' The AUC is 0.996 which corresponds to a signal to noise ratio of 3.75 indicating good correspondence between the two processing methods and also indicating that the smoothing and blocking by the JPEG compression is not strong enough to interfere with the meat texture.

Figure 5 shows two average complex cell response verses angle curves for different meat slices. Both graphs have a line for the finest resolution exam-

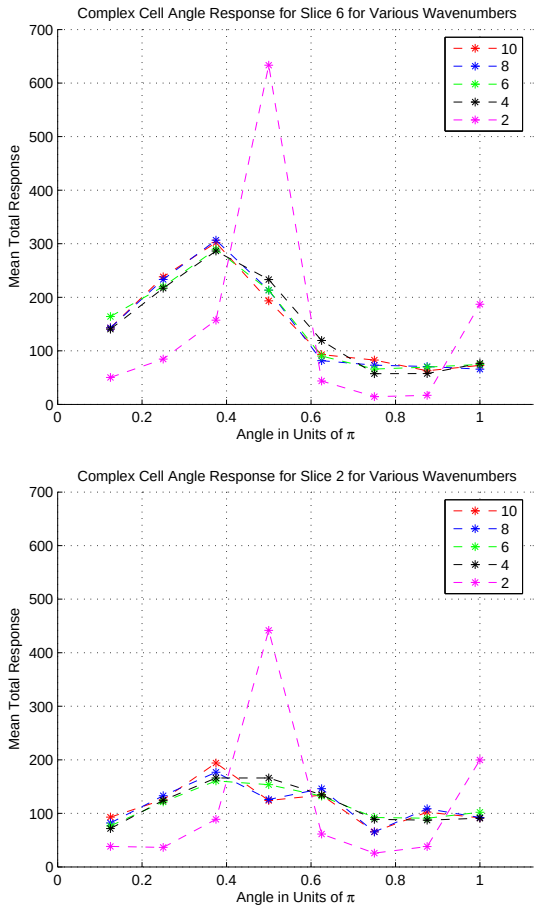

Figure 5: Complex cell responses for several wavenumbers as indicated by the legend over two meat slices. (Top) The response for a well cut slice with a clear peak below 0.4π and (below) the response from a poorly cut slice.

ined ($\lambda/2\pi$ = 2pixels) strongly peaking at $\theta = 0.5\pi$, which occurred for all slices. This is believed to be an anomalous result caused by the wavelet wavelength being too small in relation to pixel resolution. For the top graph of figure 5 the peak is just below $\theta = 0.4\pi$ from vertical, matching well the texture seen in visual inspection of the images. The bottom graph shows the peak shifted closer to horizontal which is a result of the incorrect slicing technique producing a subtle horizontal textural effect.

4 Conclusions and Future Directions

A strong relationship between colour segmentation and texture segmentation using singular values has been demonstrated in a set of meat images. Concordantly JPEG error in our images is shown to be of no consequence as it does not appear a confounding factor in the singular value texture analysis. Complex cell analysis was applied to investigate the determination of textural direction by computer vision which showed good correspondence with human visual inspection, but also highlighted how incorrect slicing of meat causes confounding texture.

Future work may involve correlation of texture analysis to NIR spectra, more consideration into appropriate texture analysis methods and 'image' processing of hyperspectral images obtained by manually sequentially scanning locations on the meat face with NIR. The primary goal is to establish, by image analysis, the localised meat content for NIR investigations. The ultimate goal is to obtain hyperspectral VIS/NIR images of sufficient spatial resolution (just what resolution is sufficient is an open question) and to consider the hybridisation of spectral analysis and image processing techniques.

References

[1] T. Pham and L. van Vliet, "Blocking artifacts removal by a hybrid filter method," in *Proceedings of the eleventh annual conference of the Advanced School for Computing and Imaging*, (Heijen, Netherlands), pp. 372–377, June 2005.

[2] Z. Wang, H. Sheikh, and A. Bovik, "No-reference perceptual quality assessment of JPEG compressed images," in *International Conference on Image Processing*, vol. 1, pp. 477–480, 2002.

[3] G. S. Gupta, D. Bailey, and C. Messon, "A new colour space for efficient and robust segmentation," in *Image and Vision Computing New Zealand*, (Christchurch, New Zealand), November 2004.

[4] W. K. Pratt, *Digital Image Processing*. New York, NY, USA: Wiley-Interscience, second ed., 2001.

[5] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. New York, NY, USA: Wiley, 2 ed., 2000.

[6] H. Barrett, C. Abbey, and E. Clarkson, "Objective assessment of image quality. iii. roc metrics, ideal observers, and likelihood-generating functions," *Journal of the Optical Society of America A*, vol. 15, no. 6, pp. 1520–1535, 1998.

[7] N. Petkov and P. Kruizinga, "Computational models of visual neurons specialised in the detection of periodic and aperiodic oriented visual stimuli: bar and grating cells," *Biological Cybernetics*, vol. 76, no. 2, pp. 83–96, 1997.