Colour and Texture Analysis for Automated Sorting of Eviscera.

Richard J. Gibbons†
David J. Williams‡

† CSIRO Meat Research Lab., Brisbane, Australia
‡ Loughborough University, UK

Abstract

This paper deals with some techniques used to recognise bovine products by use of image processing, and in particular the location of a liver within a pile of eviscera by use of a combination of colour and texture analysis.

1. Introduction

The potential to use robotics and automation in the food sector is limited by the non uniform nature of products. For example, the Automated Slaughter Project developed by CSIRO Meat Research Laboratory has had to deal with bovine carcasses varying in weight from 220 to 800kg, and the associated size differences.[1] Standard automation techniques exist for some tasks where the food product can be presented uniformly to a machine. In a slaughterhouse many handling and cutting tasks operate on unordered and variable product, to automate these sensing must be used.

Both colour and texture attributes can be used to recognise objects within a scene by matching to a set of values previously generated from test images. In organic materials often neither method can be used effectively as the target object parameters vary too much to adequately exclude other objects. Self tuning can be attempted to improve upon a particular model matching algorithm.[2] However, there is always a chance of misclassification which can be reduced if a combination of methods is used.

2. Task Description

In the Advanced Slaughter Technology system eviscera is ejected automatically from the carcass and falls upon a moving stainless steel conveying system 2 meters wide. This eviscera then has to be sorted and separated prior to inspection. To automate this the various components within the eviscera have to be identified so that they can be automatically handled and separated. The eviscera pile contains all the internal organs of the animal. (Figure 1a.) The liver was selected as the first organ to identify. This was chosen as it is the highest value product within the eviscera and is relatively simple to mechanically extract. It is often severely obscured so that little or no shape classification techniques can be used.

3. Colour Analysis

Food products in general, and specifically animal products are not easily identifiable by colour. This is because, unlike humans, computer imaging techniques are not specifically tuned to respond to the part of the colour spectrum which most the food components' colours lie in. An additional problem encountered within abattoir operating conditions is that the ambient lighting used is often sodium which does not radiate many spectrum lines. For this test it was chosen to augment the background lighting with more sodium sources rather than confuse the conditions by adding light of a different spectrum. This has the effect of causing a fairly uniform lighting density

BMVC 1991 doi:10.5244/C.5.42

over the viewing area and any light spill from the existing sodium sources will not alter the colour characteristics of the objects.

The Hue, Saturation, Intensity (HSI) colour space was used for analysis. This scheme has advantages over the conventional Red, Green, Blue (RGB) system as intensity is already isolated from the other components of the image so algorithms can easily be tuned to isolate colour components of the image.

For the colour analysis all three components of the image were used and a target set of HSI values for the liver arrived at by inspection. The difference, δ, between these values i_t, s_t, h_t and observed values, i, s, h were calculated and summed. (This gives the D_4[3] or 'city block' distance from the observed and target values.) However the hue value of a pixel lies on a cyclic scale, as hue is an angular measure in colour space. The difference between two hues has to be adjusted to allow for this.

$$\delta = |i_t - i| + |s_t - s| + min(|h_t - h|, h_{max} - |h_t - h|)$$

It was felt that no significant advantage could be gained from calculating the Euclidean or absolute distance between observed and target values. This is because a target point is often smeared along one of the axis of the HSI co-ordinate system and the Euclidean distance has the effect of favouring small deviations in each of the co-ordinate planes rather than one larger deviation in one plane. However the primary reason for using the D_4 distance was to save computational overhead. The resultant output of this classification operating on figure 1a is shown in figure 1b.

4. Texture Classification

Texture analysis uses the relationship between neighbouring pixels' grey levels. Over a area or patch the relationship can be statistically summed in various ways to reach a value of texture for that patch. This can be used to split images into regions with distinct boundaries. However if the objective is to classify individual pixels, a patch method cannot be used. To replicate the data required for statistical analysis which is usually provided by the patch a texture measure is found for a pixel over several directions in the image. This is acceptable for amorphous food objects under uniform lighting conditions as there should be no directionality to the texture of an image.

For each pixel in the image, therefore, it is possible to get a texture measure by considering the absolute difference between the intensity value of that pixel from others at separation d from it. This is done in all directions to form a box area around the pixel, of side length $2d+1$ or those pixels with D_8[3] distance equal to d. This can theoretically give a difference between $8di_{max}$ but in most cases will be much less. This measure, T, gives a value to the average rate of change in intensity of the image at any point in $8d$ directions. As d is increased smaller frequencies will not be detected and only lower frequency changes will be recognised but boundary effects of the objects come into play. As the analysis is averaged over several directions any overall gradient changes (i.e. from lighter to darker) cancel out leaving only the texture gradient measures.

By experimentation a value for d of 5 gave a reasonable measure of texture for the images, resulting in a low score for smooth surfaces such as the liver and background steel tray and high scores for grainy areas such as the lungs, blood pools and heart. The resultant output of this classification operating on figure 1a is shown in figure 1c.

5. Combining the methods

Each individual method misclassifies some pixels either by identifying them as liver incorrectly or not classifying as liver. For this case it is more important that that which is not liver is not identified as liver. As the primary function of the analysis is to pass information to a robotic system adequate enough for location of the liver, the combining method is aimed at reducing the misclassifications of non-liver as liver.

For an image comprising of 256 grey levels (or 8 bits) $\times$ 3 planes, δ will lie in the range of 0 to 640. To isolate pixels which are potentially part of the liver this range is thresholded, such that values greater than the threshold were deemed not to be liver. Although this excludes 'non-liver' pixels it includes others which cannot be eliminated by further colour classification. However when the texture rating, T, of these selected pixels is investigated it is found that the intercept of low T scores and low δ is enough to identify a significant percentage of the liver and to only classify few non-liver pixels as liver. This method yields a frame, the intensity of each pixel (mainly zero) representing the closeness that it matches the liver characteristics. The resultant output of this method of combining the two classification methods operating on figure 1a is shown in figure 1d.

To convert this frame into a set of co-ordinates for the position of the liver the median value of the frame is used. This was chosen as an easy way to interpret this information to gain some measure of accuracy of locating the liver.

6. Results

100 previously captured images of eviscera were used to test the analysis. These images are 512 $\times$ 384 pixels and 3 $\times$ 8 bits deep. The position of the liver in the eviscera was ascertained and recorded manually so that the distance from actual to calculated could be found and this was attempted to be minimised. At first some manual tuning was used to arrive at rough values for i_t, s_t, h_t and the threshold. (d already fixed at 5.)

From the rough initial values the analysis was run on the 100 samples with small perturbations to the values to 'tune' the system to the optimal value set. For the tuning the Euclidean distance from calculated liver centre to actual liver centre was summed for all 100 samples. This was minimised. After a set of values was reached for i_t, s_t, h_t and the threshold, the 100 samples were inspected for the individual Euclidean distances from calculated to actual to confirm that these distances were mainly small (<50 pixel units) indicating a hit in the liver area rather than many near misses which would yield the same effect.

For the perturbation step of five (i.e. the minimum change of the target values was five) the optimal point was found at i_t=35, s_t=185, h_t=185 and the threshold = 70. It was found that 68 samples were within the 50 pixel limit initially indicating a found liver in 68% of the cases. After closer inspection of this optimum case that 65 livers were correctly found, 12 were close to the liver area but missed the absolute boundary while 23 livers were missed completely.

7. Discussion

Although this algorithm doesn't find all the livers this is mainly due to the crude method of taking the weighted potentially liver pixels and using their median as the position of the liver. In fact the colour and texture analysis tends to find distinct areas of isolated pixel clusters. These clusters relate to areas within the image that are close to liver in properties and are blood clots, lungs filled with blood, etc. Usually the area within the output frame that represents the liver contains the largest amount of pixels and these pixels are of the greatest intensity. However the frequency of the smaller clusters distinct from the liver are enough to occasionally move the median out of the liver's area. This is seen as a shortcoming in the interpretation of the reduced data set produced by the pattern classification rather than the method itself.

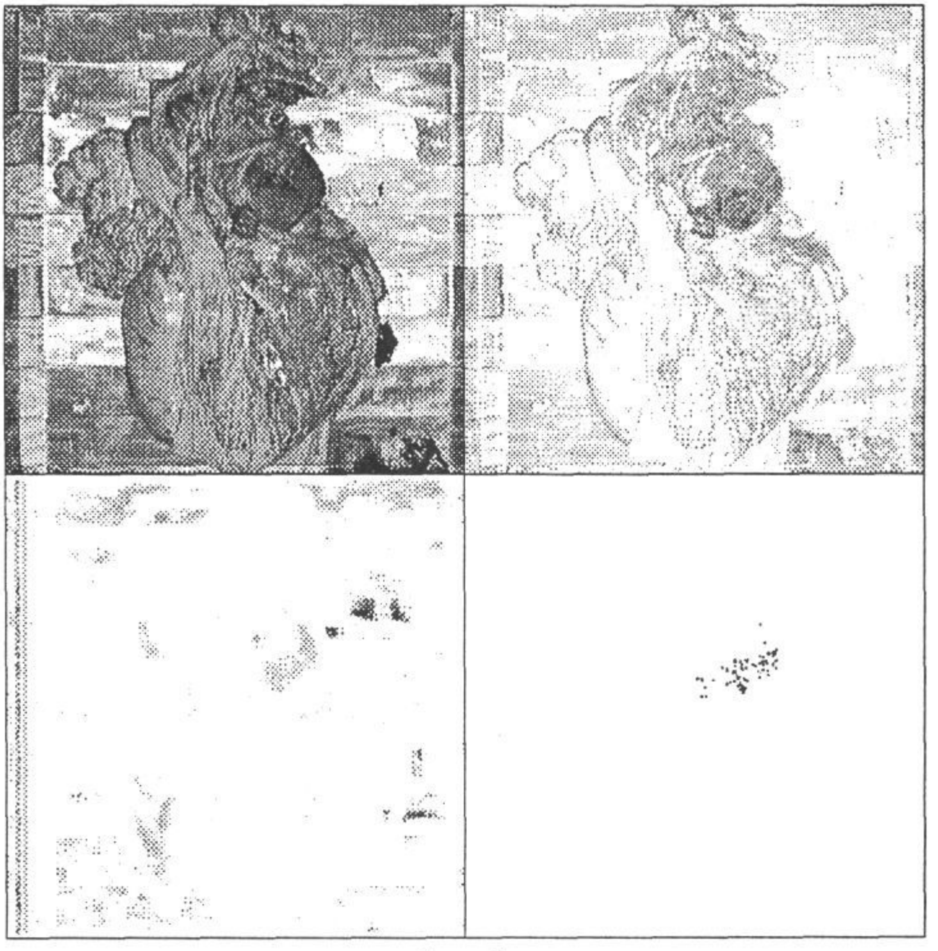

Figure 1
(a) Video image of eviscera (b) Colour analysis output
(c) Texture analysis output (d) The two methods combined (highlighted)
N.B. The liver is a little to the right and above the centre of the image.

References

1. Buhot J. W. & Rankin R. *Automated Slaughter and Dressing of Bovines* IARP Workship on Robotics in Agriculture and the Food Industry, Avignon 1990, pp 299-303.

2. Gibbons R., Williams D. J. & Green D. A., *Heuristical Methods of Algorithm Selection for the Orientation of Diverse Non-uniform Products* 3rd National Conference on Robotics, Melbourne, June 1990, pp 282-292.

3. Gonzalez R. C. & Wintz P., *Digital Image Processing* Addison Wesley, 1987.