

More machine vision applications in the NCEA

John Billingsley

National Centre for Engineering in Agriculture
University of Southern Queensland
Toowoomba, QLD AUSTRALIA

1. Introduction

In the early nineties, the research team at the National Centre for Engineering in Agriculture established a reputation for vision-based automated guidance of agricultural vehicles.[4] This work has a new lease of life with recent funding. A succession of further vision projects have been somewhat unusual, ranging from the visual identification of animal species for the culling of feral pigs to vision-based counting of macadamia nuts. A unifying feature is the easy availability of low-cost cameras and a framework for integrating analysis software using DirectX 'filters'. Machine vision has changed from its earlier status as a sophisticated and expensive technology to a low-cost solution for more general instrumentation.

For the rapid solution of ad hoc problems, it is easy to exploit the convergence between computing and media entertainment. A system including camera, interfacing, gigabytes of data storage, display and an embedded PC card can cost less than a single conventional high resolution industrial camera and interface. However we are also pursuing a more fundamental approach, building systems around image-sensor components interfaced by means of RISC processors.

2. Identification of Animal Species

In the Australian Great Artesian Basin, there has been a programme of capping bores and piping the outflow to watering points.[13] Access to water can therefore be controlled for feral and native animals, in addition to the farmed livestock. The objective is to allow normal access to both farmed and protected species, while 'undesirables' such as feral pigs are directed to a second water supply in an enclosure from which they cannot escape until they are 'harvested'.[8] Feral pigs do hundreds of millions of dollars worth of damage per year, but there is a lucrative export market for wild pork.

Shortly, all Australian farmed animals will carry tags under the National Livestock Identification Scheme. This would make the task simple were it not for the

need to allow kangaroos, emus and some other wild species equal access to water. The NLIS tag will give a reassuring double check, but the task of species identification remains.

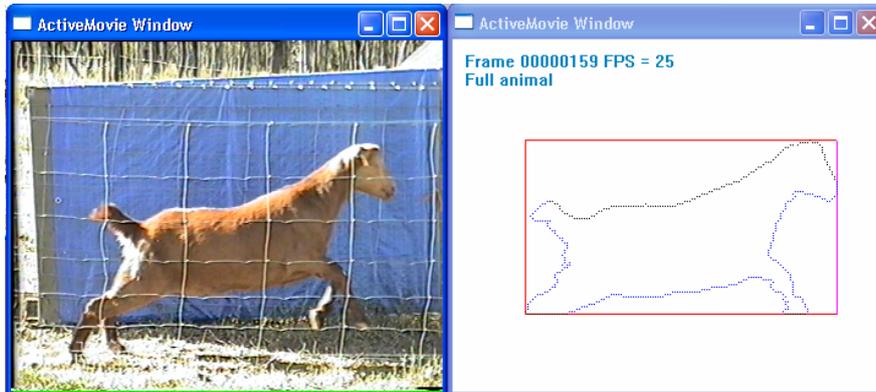


Fig 1. Goat and boundary

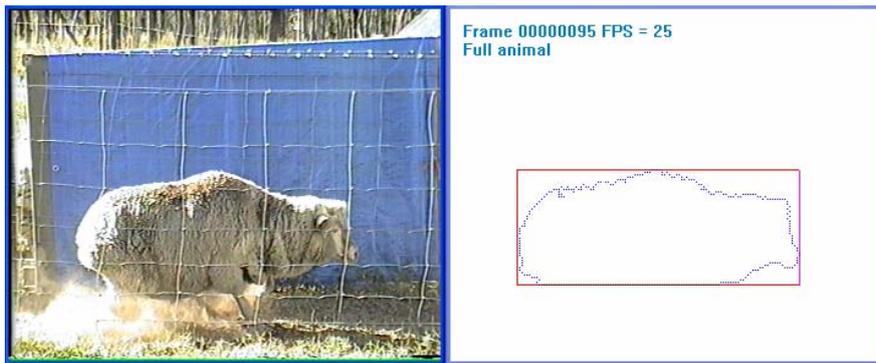


Fig 2. Sheep and boundary

The animals must approach the water through a narrow fence-wire corridor, at the end of which is a gate which is controlled to direct them into one or other of two compounds. One of these has a one-way exit while the other is closed. Initial experiments used a blue tarpaulin behind corridor, to give easy colour-based discrimination.

Identification of the animals is based on their boundary shape. Many edge detection methods lose important information by locating a boundary as a scatter of points with no regard to the order in which they should be linked to form a curve that circumscribes the shape. We therefore use a chain method that 'draws' the boundary, while a convention of 'animal is to the left of the boundary' defines the direction of drawing to be anticlockwise.

A search is first performed on a coarse grid until the first 'animal' pixel is detected, then adjacent pixels are searched upwards to find the boundary. The boundary is then traced, using a strategy first proposed in 1983. [2, 3]

Points are 'stitched' around the boundary in a routine that can most easily be likened to a dance step. Assume that we are 'dancing' to the left, starting from a point that has been identified as 'animal'. We step forward and test the new point. If it has changed to 'not animal' we note the midpoint as a boundary point and step diagonally backwards to the left. If on the other hand it is 'animal' again, we rotate our body and progress direction forty-five degrees to the right and take a further step forward. A Javascript demonstration of the full algorithm can be found at <http://www.jollies.com/stitch.htm>.

Points are found in a sequence that traces the boundary of the animal. We obtain a sequence of vectors of the nature of a Freeman chain. These give vectors, each of which takes us further around the boundary and has a direction, ψ . When we plot ψ against the boundary distance s we obtain an 's- ψ ' curve relating tangent direction to the distance advanced around the boundary. A complex two-dimensional image has been reduced to a simple one-dimensional function that can be matched against a set of shape templates of the same form.

If the object is rotated, a constant is added to all the points of the curve. The distance can optionally be normalised to allow shapes of different sizes to be matched. Matching is simplified if we can choose a unique starting point on both curve and template.

Instead of the full circumference of an animal, the 'top half' is usually sufficient for identification and avoids the confusion that leg movements can give. The edge is traced from the upwards vertical tangent at the nose, along the back to the downwards vertical tangent near the tail.

It was felt that the blue tarpaulin might deter many of the animals in the wild from approaching the system. Further experiments discriminated between animals and background on the basis of the difference between the frame and a 'remembered' image of the background. It was harder to get a complete circumference this way, but a 'bounding box' was easy to construct. An elementary algorithm based on the two arrays of distances from the box to the front and to the rear of the animal gave good discrimination.



Fig 3. Natural background and bounding box.

The project is at a field testing stage (literally) and sheep and goats can be separated with 100% success, except when animals overlap seriously.

3. The Counting of Macadamia Nuts

One of the objectives of breeding varietal strains is to produce a tree with maximum yield. The harvesting method is simple. Nuts fall from the trees and are picked up from the ground. The trees are planted close together, however, so the task of segregating the catch from the individual 'drop zones' is important. [1]

The accepted harvester technology consists of a 'bristle roller' several metres wide. The nuts, in their husks, are trapped between the bristles and carried upwards, where they are stripped from the roller by 'fingers'. They fall into an auger that carries them across the machine to another auger at the side, which in turn carries them to a bin at the back of the machine.

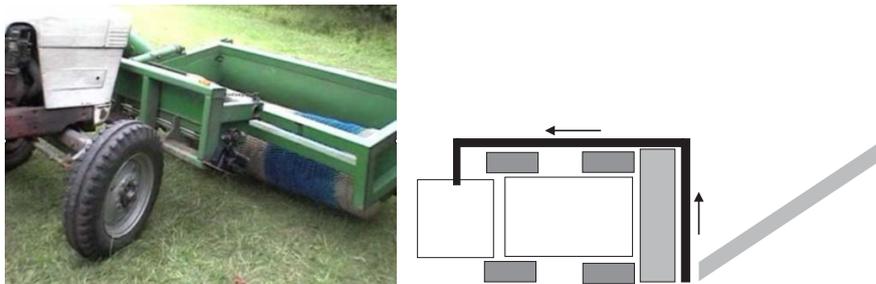


Fig 4. Harvester photo, diagram

As they move to the left in the auger at the front of the machine, nuts that have been gathered from the right-hand-side of the swathe are joined progressively further to the left. At any instant, a point in the delivery auger will thus contain nuts gathered from a diagonal stripe of the width of the roller. If this were the basis of a count, it would be impossible to deconvolve the 'Green's function' to ascribe the count rate to a fine enough location to assess the yield of individual trees.

Machine vision has been adopted as the solution. The bristle roller has been coloured blue, to obtain good colour contrast with nuts at all stages of maturing, and cameras have been mounted inside the housing.

First the RGB pixel components must be analysed to determine 'roller' or 'not roller'. It was found by experiment that $(red > blue)$ gave good discrimination. Unfortunately leaves are also picked up that are of a similar colour to the nuts. The second stage must therefore involve shape discrimination to select the circular nuts and reject other shapes. Figure 3 shows the discriminated image when all pixels are processed, but the chosen algorithm requires only a small proportion of the pixels to be sampled.



Fig 5. Camera locations

The formal Hough transform approach would involve applying a filter to the discriminated data to identify all boundary points. Triplets of boundary points can then contribute 'votes' on the circle centres.

Because analysis must be performed in real time, a much simpler circle-detection algorithm was chosen. Pixels are examined on a coarse grid, with spacing half the size of a nut. The grid is such that columns of points are examined downwards, the columns progressing from left to right. When a pixel is found that is 'not blue', adjacent pixels are then explored from this point up and down to find two points on the boundary. From the midpoint of this chord, pixels are explored horizontally to find boundary points; the centre is located and the radius is determined.

To check the circularity of the object that has been found, we inspect points on the vertices of two octagons. These lie on two circles, one inside and one outside the circle we are testing.



Figure 6. Detected nuts and leaves

Although rough and ready, this algorithm is robust and rapid. If the initial search grid has intervals of five pixels, only four percent of the pixels are initially inspected. For each nut or leaf detected, some further thirty to forty further pixels are tested.

Having determined the coordinates of nut-centres in one frame of the video stream, it is necessary to collate the sequence of frames to ensure that nuts are neither counted twice nor omitted. The result is a map of coordinates relative to the location of the tractor.

Of course that is only half of the story. In order to aggregate the totals tree by tree, we must determine the location of the tractor relative to each tree.

Machine vision provides a solution once again. At first glance, it would appear that precision GPS satellite navigation could give the tractor location. In practice, the canopy of trees will degrade the accuracy to a point where it cannot be used. Instead, a side-looking camera observes the tree trunks as they stream past some carrying an identifying marker so that the count cannot lose step. Trees in more distant rows are recognised by their 'streaming rate', so that they do not confuse the issue.

4. Animal Behaviour

There is a breeding programme to improve the survival chances of a species of dunnart, *Sminthopsis douglasi*. A serious hazard is the aggressive nature of the small mouse-like marsupials. If the female is not in oestrus when introduced into a cage with the male, there is a danger that they will fight to the death.

A first step was to mount a camera where it could capture video from two adjoining cages containing male and female. Background discrimination enabled the animals to be located as they moved, easing the task of the student who had to monitor the recorded video to judge when it would be safe to put the animals together.

The next step will be to encode the movement to obtain an automatic assessment of the animals' behaviour. Initially this will merely alert the animal breeders, but eventually automatic operation of a gateway between the cages might be possible.



Fig 7. Dunnarts and detected movement

5. Texture Analysis

One of the factors used to determine citrus quality is the texture of the skin.

The skin texture of citrus fruit is a combination of three different types of spatial variation. Sub-millimetre wrinkles cover the entire skin and are irregular, but have relatively constant coverage. Small dimples 1mm-5mm in depth are randomly spaced around the fruit. It is the depth and quantity of these which have the greatest impact on the skin texture grade. The third type of variation is deformation from the normal spheroid shape. These lumps or flat spots can be caused by rough handling, or may be due to variety.

The texture can be measured directly from a fruit using an expensive stylus instrument where, similar to that of a record player, a needle touches the skin of the fruit as it revolves. [11] The changes in position are amplified and recorded. A serious problem is that this method only gives a single sample from one 'latitude' around the fruit, which may or may not be representative.

The machine-vision solution is to illuminate the fruit from the side, so that to the camera mounted in front of it, it appears as a 'half moon'. [7] The 'terminator', dividing lit and portions in shadow, will appear as a ragged vertical line, with a statistical distribution of the horizontal 'roughness' that is readily related to the texture. As the fruit rotates, a sequence of measures can be accumulated to give an assessment of the entire fruit.



Fig 8. Orange, illumination and camera

One such measure involves detecting the shadow edge position as a function $x(y)$ of the scan line y on which it is seen. This function can be doubly low-pass filtered by executing code equivalent to

```
xsmooth(ymin) = x(ymin)
for y = ymin to ymax
    xs = xs + (x(y) - xs) / k
    xsmooth(y) = xs
next
for y = ymax to ymin step -1
    xs = xs + (xsmooth(y) - xs) / k
    xsmooth(y) = xs
```

next

where k is a smoothing parameter, a sort of 'distance constant'.

The high pass 'roughness' signal is given by $x(y) - x_{\text{smooth}}(y)$. A measure of texture can be obtained by squaring and summing this. We can tune k to remove lumps and deformations, while preserving dimples.

A second project involving vegetation texture analysis is not really agricultural. There is a need for a fast and simple way of analysing the cover of a football pitch, to ensure that there is no bald and slippery patch where an expensive player might suffer damage. [9]

6. Measuring the Density of Dingo Teeth

A biologist colleague had a requirement to measure the density of porous dingo teeth, in order to establish the validity of using the relationship between tooth density and age to estimate the age of the animal.[5] Skulls containing canine teeth from 68 'known-age' animals, either field-captured or captive-reared, were borrowed from dingo skulls held by CSIRO Sustainable Ecosystems, Canberra, Australia.

Weighing the teeth was of course not a problem, but their very porosity was at variance with the use of an 'Archimedes' immersion method for measuring their volumes. The biologist seized on the suggestion that machine vision could be the answer.

The original intention was to use 'structured light' to map the tooth. Then there was a surprise announcement that the museum required the return of the teeth by the end of the week. Attention turned to a means of capturing the data for later analysis. The result was almost certainly superior to the original intention.

The canine tooth, shaped rather like a banana, was attached to the vertical axis of a small stepper motor. The motor had been 'recovered' from a discarded floppy-disk drive. A 'Smartcam Pro' web camera was mounted firmly to observe it. It was decided that the 320 by 240 pixel resolution would be best employed by mounting the camera in 'portrait' position. The off-white tooth was illuminated in front of a black background.

Only 50 images were to be captured per revolution. Even so, the prospect of saving and later processing over eleven megabytes of data for each of 160 teeth was daunting. So just the green signal was captured, yielding a clear binary silhouette. The data was further reduced before saving.

For each of the 320 'slices' of the image, the location of maximum and minimum tooth boundary were found. If no white appeared because the slice was beyond the end of the tooth, two values of 0 were recorded, otherwise two comma-separated numbers were written to file. The file size for each tooth was reduced to just over 100 kilobytes.

Some industrious work by a biology student saw the teeth scanned and returned to the museum on time. The problem still remained of processing the data.

Each of the number-pairs in the data file represents a line section of the perceived tooth image, in other words a 'left tangent' and a 'right tangent' to the tooth for a given elevation angle from the camera lens. The 'perspective effect' can readily be accommodated by drawing the tangents through a single point representing the camera.

The method adopted was similar to that of the tomograph. A planar array of points was set up, initially deemed to be 'occupied'. Now each tangent becomes a line that sets a boundary between occupied and unoccupied points and those that are unoccupied are eliminated. For the next image in the sequence, the camera position is rotated 1/50 revolution with respect to the plane and the process repeated. At the conclusion, the survivors are counted to give the area of the slice.

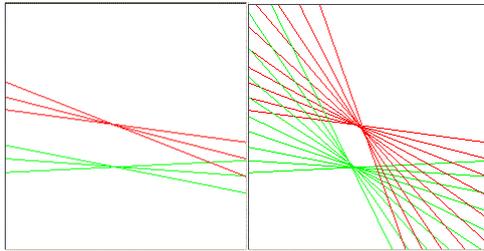


Fig 9. Building image from left and right tangents

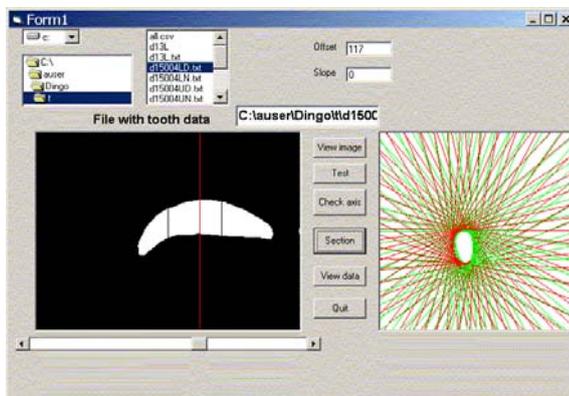


Fig 10. Left and right estimates of the slice.

The left and right tangents give two distinct estimates of the area. By comparing them, corrections can be made for errors in the location and angle of the image of the axis. Totalling the area gives a measure of the volume.

7. Vision guidance

The original vision guidance method had a number of patented features. These have been carried through into the new version of the system.

The basis of the strategy is to form 'keyholes' in the image, each of which embraces a single row. A regression fit is made to fit a line to the 'plant' pixels in the keyhole, updating the estimate of the row's location and direction. This in turn updates the keyhole to be used for the next frame.

The vegetation or furrow that must be followed must be discriminated from the background, either by comparing brightness or colour components. Since the lighting is apt to change, this could give problems. To solve this, the threshold level is manipulated to give a certain proportion or 'plant' pixels in the keyhole. At any stage of growth, this will be fairly constant. Figure 11 shows a typical image from the row-following algorithm.



Fig 11. Fitting lines to crop rows

8. Conclusions

A vision approach can be applied to an ever broader range of instrumentation tasks. It has become simple to exploit the media-motivated interfacing and video-stream processing tools that are now readily available.

Projects not described here include a refractometer for measuring sugar cane juice, based on a line-scan camera.[12] The interfacing task would have been simplified if a cannibalised webcam had been used instead. There is a divergence within computer peripheral systems. Analogue to digital converters for instrumentation have become increasingly complicated and costly, while sound cards with high performance ADCs are available at give-away prices. Transducers for

absolute position are extremely costly and we have given serious thought to using a webcam to inspect a measuring tape. It may be technological overkill, but it is a low-cost alternative to a commercial sensor. There will be many occasions on which it is easier to tailor an elaborate consumer product rather than craft a simpler engineering solution.

The projects presented in this paper may seem strange, but future applications may well push the bounds of imagination much further.

Acknowledgments

This review draws heavily from the work of Mark Dunn, Kerry Withers and others in the NCEA and the University of Southern Queensland.

References

- [1] J Billingsley, "The Counting of Macadamia Nuts", *Mechatronics and Machine Vision 2002: Current Practice*, Research Studies Press Ltd. ISBN 0-86380-278-8, pp 221-7.
- [2] J Billingsley and A A Hosseinmardi, "A low-cost vision system for robotics", *Proc. British Association meeting*, Brighton, August 1983.
- [3] J Billingsley and A A Hosseinmardi, "Low cost vision for robots – a pragmatic approach", *IEE Colloquium (Digest)*, n 1983/65, 1983, p 3. 1-3. 2
- [4] J Billingsley and M Schoenfish, "Vision and Mechatronics Applications at the NCEA", Fourth IARP workshop on Robotics in Agriculture and the Food Industry, Toulouse.
- [5] J Billingsley and K Withers, "Measuring the density of dingo teeth with machine vision", *Sensor Review*, v 24, n 4, 2004, p 361-363
- [6] E R Davies, *Machine Vision Theory Algorithms, Practicalities* (3rd ed.). Elsevier.
- [7] M Dunn and J Billingsley, "A Machine Vision System for Surface Texture Measurements of Citrus", *Proceedings 11th IEEE conference on Mechatronics and Machine Vision in Practice*, Macau, November 2004, pp 73-76.
- [8] M Dunn, J Billingsley and N Finch, "Machine Vision Classification of Animals", *Mechatronics and Machine Vision 2003:Future trends*, pub Research Studies Press Ltd, Baldock, UK, ISBN 0 86380 290 7, pp 157-163
- [9] M Dunn, J Billingsley, S Raine and A Piper, "Using Machine Vision for Objective Evaluation of Ground Cover on Sporting Fields", *Proceedings 11th IEEE conference on Mechatronics and Machine Vision in Practice*, Macau, November 2004. pp 88-91.
- [10] V Kindratenko, "On Using Functions to Describe the Shape", *Journal of Mathematical Imaging and Vision*, vol. 18, no. 3, pp. 225-45.
- [11] R Leach, *The Measurement of Surface Texture using Stylus Instruments*. 2001, National Physics Laboratory: London. p. 79.
- [12] S McCarthy and J Billingsley, "A sensor for the sugar cane harvester topper", *Sensor Review*, v 22, n 3, 2002, p 242-246
- [13] J Seccombe, "Sustainability of the Great Artesian Basin", *Proc. of Academy Symposium*, Canberra, Australia.

