Proceedings of the 2004 New Zealand National Conference on Non Destructive Testing
June 27-29, 2004, Palmerston North, New Zealand

HIGH SPEED WEIGHT ESTIMATION BY IMAGE ANALYSIS

D.G. Bailey*, K.A. Mercer*, C. Plaw*, R. Ball**, H. Barraclough**

*Institute of Information Sciences and Technology
Massey University

**Institute of Technology and Engineering
Massey University

Abstract

When produce is prepackaged and sold by weigh it is often necessary to weigh each individual item. Typical packing lines operate at 10-15 items per second making conventional weighing methods inaccurate. Image analysis techniques are employed to overcome these problems by measuring the volume of each item in two perpendicular views. A system of mirrors is used to obtain the views within a single image. The system achieves an accuracy of 4.4% per item or 2% per 100 g bundle at rates of up to 30 items per second.

1. INTRODUCTION

Most produce – fruit and vegetables – is sold by weight. When the produce is prepackaged, it is important that all of the items within a package are approximately uniform in size and quality. If the package weight is arbitrary, this is relatively easy to achieve, but for fixed weight packages there is the added complication of ensuring that the stated weight is provided. Economic considerations require that the package weight is not significantly exceeded, effectively giving away produce. When each package contains a large number of items, and the packages can be weighed while packing, this criterion can be achieved. However, when the package weight is small, with each package containing only a few items, the variability in package weight can be significant. Maintaining uniform item size within a package results in the weights being quantized, and guaranteeing that the minimum weight is provided without significantly exceeding the stated weight can be very difficult.

A second issue with regard to grading and packing is that such systems have a high throughput. Grading of the produce is a bottleneck in many packing houses, therefore it is desired minimize the time spent on each individual item. The settling time of the load cells, vibration and bouncing of the produce as it settles on the load cell provide an upper bound to the throughput of the complete grading system. With a target throughput of 15-20 items per second, physically weighing the individual items is very expensive and impractical. To overcome these problems, size based grading is often used as a substitute for weighing each item of produce.

Grading on size will only produce acceptable results if the volume can be accurately estimated from the visual measurements made. It also requires that the density of the produce is constant, or at least consistent within a batch. Without these two conditions, the items within a package may well be of uniform size, but no claims can be made about the weight of each package. As the variation in density increases, and the accuracy of the volume estimation decreases, the average package weight must increase in order to ensure that the minimum package weight is exceeded.

For items that are approximately spherical and are not easily bruised, mechanical screens provide an effective and fast method of size grading. Each screen usually consists of a mesh with certain hole size. The mesh is vibrated, and the items that are smaller than the holes fall through the mesh, while those larger than the mesh size remain on top. By having a series of meshes, the produce can be sorted into several size categories.

This approach is generally unsuitable for produce because it relies on mechanical jostling to get the produce to fall through the mesh, and this may bruise or damage some items. It is also unsuitable for items that are long and thin, because the size of each item then depends significantly on its orientation.

An alternative approach is to make one or more measurements of the item being graded using machine vision. Such measurements are usually based on a projection, either of the whole object or part of the object. Projection based measurements require that the object can be segmented from the background. This can be achieved by placing each item against an appropriate background, and using suitable lighting. Any projection based method implicitly makes assumptions about the shape of the object, but can be effective if the shapes are sufficiently consistent that these assumptions are valid. If only a part of the object is measured, it is assumed that the rest of the object has similar characteristics, or can at least be modeled and predicted from the measurement made.

2. APPLICATION DETAILS

In the application considered, the individual items of produce varied in weight between 5 and 50 grams. These items are to be packaged into bundles with a total weight of 100 grams. Each bundle should contain between three and seven similarly sized items, with the very large and very small items rejected. Such a wide weight range per item, and a small item count per bundle, requires that each item be weighed so that it may be allocated to an appropriate bundle. The input weight distribution is also continuous, with individual items not having convenient discrete weights for allocation directly to bundles.

The produce has approximately uniform density that also varies little from one item to another. Therefore measuring (or estimating) the volume provides a good estimate of the weight of the produce. This was verified by measuring both the weight and volume of a large number of items (the volume was measured by water displacement). From the volume alone, it was possible to estimate to weight to within 1% accuracy. Therefore estimating the volume of each item provides an alternative to actually weighing the items.

As each item is approximately cylindrical, the most obvious model is to assume a circular cross-section model each item as a generalised cylinder. The volume may be estimated from the projection of the item by measuring the diameter, D, as a function of the length (as illustrated in figure 1). The volume is then given as

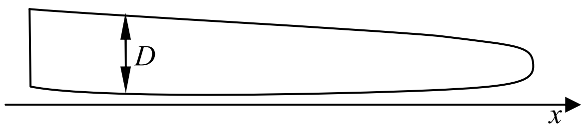

Figure 1: The projection of an item. Measuring the diameter as a function of length allows the volume to be estimated.

$$\begin{aligned} V &= \int A(x)dx \\ &= \frac{\pi}{4}\int D^2(x)dx \end{aligned} \tag{1}$$

2.1 Physical Setup

To facilitate image capture, the produce is singulated onto V shaped cups. With an appropriate trigger arrangement, this allows an image to be captured of each item of produce. The camera has the horizontal axis aligned with the produce to simplify the processing. The diameter can be measured by counting the pixels in each column of the image within the produce, D_x. Equation (1) can then be discretised as:

$$V \propto \sum_x D_x^2 \tag{2}$$

where the constant of proportionality depends in the size of each pixel and is found by calibration. This approach of estimating the volume from a single projection did not give accurate results. There was sufficient deviation from a circular cross-section that this assumption limited the accuracy to about 10%.

To estimate the volume more accurately, it was obvious that the eccentricity of the cross section of each item also needed to be measured. One possibility that was investigated was to image the end of each item, and assume that the cross section scaled along the length with the visible diameter. This improved the accuracy of the volume estimation to about 6%. However, the accuracy could not be increased further because of the limited resolution in measuring both the end cross-section, and the diameter at each point along the item.

Another possibility that was tried was to measure the projection of two views taken 90 degrees apart. Figure 2(a) shows the measurement of the major and minor ellipse axes from ideally placed projections, giving a cross section area

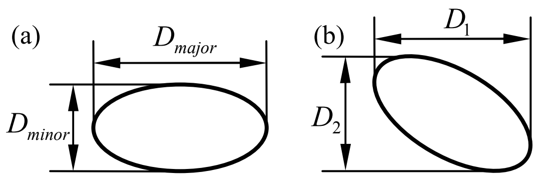

Figure 2: The perpendicular projections of an elliptical cross-section. (a) Ideal projections of the major and minor axes. (b) A more realistic set of measurements.

$$A = \frac{\pi}{4} D_{minor} D_{major} \tag{3}$$

Unfortunately, we cannot rely on perpendicular views aligned with the ellipse axes. A more realistic situation is illustrated in figure 2(b).

Use of equation (3) as

$$A \propto D_1 D_2 \tag{4}$$

results in significant error as the cross-section becomes more elliptical, and is oriented close to 45°. Better results may be obtained from

$$A \propto D_1^2 + D_2^2 \tag{5}$$

which effectively averages the volume estimates obtained from each of the two views. While this provides insufficient information to estimate the actual eccentricity, the weight estimate using this approach is within 2-3% of the actual item weight.

The two perpendicular views could be obtained from two separate cameras. However because each item is relatively long and thin, the use of a single camera was considered, with mirrors to obtain the two views. One physical arrangement that achieves this is shown in figure 3. While two views could be obtained using a single mirror, with one direct and one reflected view, this arrangement has two advantages. The first advantage is that, by symmetry, the two 90° views will have the same scale because the path length is the same. The second advantage is that it provides a third view that may also be used for grading purposes.

It was found that using the third view to also estimate the weight of the produce item actually increased the errors. This is because the central view may give significantly different results depending on the orientation of the ellipse.

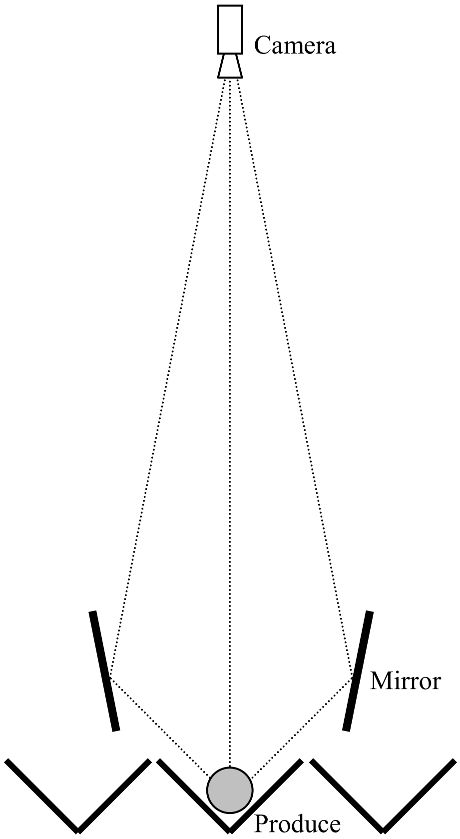

Figure 3: Two views obtained 90 degrees apart using mirrors.

One limitation of the arrangement of figure 3 is the small clearance between the bottom of the mirror and the top of the cup. If an item of produce was not completely within a cup, it occasionally caught on the bottom of the mirror and fouled the system. To achieve better clearance, the mirror needs to be lifted, and moved further from the centre-line. To fit all three views within a single image then requires a wider angle lens, with a consequent reduction in resolution of the whole image.

This may be overcome by introducing a second set of mirrors within the field of view, however the basic principle remains the same as that in figure 3. One issue with regard to using mirrors to obtain multiple simultaneous views is the longer path length of the two side views relative to the centre view. The consequence is that the camera must have a longer depth of focus to obtain well focussed images in all three views. The depth of focus may be increased by reducing the aperture, at the expense of requiring stronger illumination to maintain an adequate exposure. Such constraints impose a practical limit on what may be accomplished with mirrors.

2.2 Lighting

It is essential that the lighting be such that the item can be segmented from the background. Of particular difficulty is segmenting the produce from the cup. The ends of the produce that extend past the ends of the cup are against a black background and can easily be segmented. To aid the segmentation, black cups are used. This provides good contrast in the direct view, apart from specular reflections from the ends of the cup where it curves.

With the cups moving continuously past the camera, it is important to freeze the motion in the images. With a throughput of 15 items per second, the conveyor is moving at 75 cm/s which corresponds to approximately 2250 pixels per second in the images. Therefore to reduce the motion blur to less than one pixel, an exposure time of 1/2500 second or shorter is necessary. This rate is well within the electronic shutter capabilities of solid state cameras, so continuous illumination was used. The short exposure and small aperture (required to give adequate depth of focus) necessitate high intensity illumination. DC powered quartz halogen lamps were used to provide flicker free lighting.

The most difficult part of the image to illuminate effectively is along the sides of the items between the item and the cup. To achieve a relatively uniform light distribution a total of eight 40W lamps are used. Two lamps illuminate the top of the produce from each end. These are positioned at an angle in an attempt to provide better lighting along the sides of the produce. Within the small space, the most even illumination was obtained by directing narrow angle beams to the opposite end of the produce as illustrated in figure 4. A polished aluminium reflector directed light back along the produce. An additional two wider angle beams were used to light the produce from each side at the ends to ensure that there was adequate lighting when the produce hung over the ends of the cups.

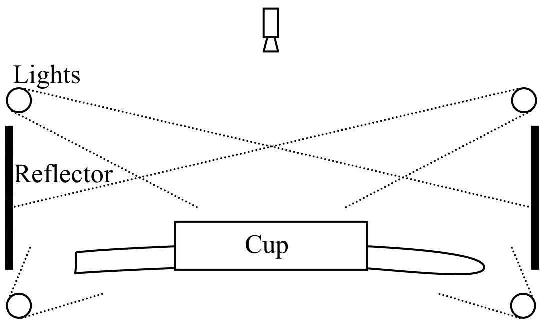

Figure 4: Lighting arrangement to give uniform light distribution over the produce.

3. IMAGE ANALYSIS

Processing speed is critical because at 15 items per second there approximately 60 ms to process each image. For this reason, after image processing algorithms were developed on a general purpose image processing system [1], they were rewritten and optimised for this particular task.

3.1 Determining if an item is present

When there is no item present in the cup, there is specular reflection from the bottom of the cup. This can makes it appear that the cup contains an item even when none is present. Since the item normally extends past the ends of the cups, the best method of determining the presence of an item is to check past the end, where the object can easily be segmented against a dark background.

The mean and variance is calculated within a 96x1 region from the central view of the item. When no item is present, only the dark background is seen so both the mean and variance will be small. When both exceed a threshold, this indicates that an item is present.

If an item of produce is detected, the procedure continues with the rest of the processing.

The first step is to assess the exposure of the image. For this a histogram is taken, and the 99th percentile determined to give a highlight value (this allows for any specular reflections from the cup). If the highlight value is greater than 250, the exposure is too long, and there is the danger of losing detail through the image saturating. If the highlight value is less than 192 there is insufficient light because the exposure is too short. If the highlight value is out of range, the electronic shutter of the camera is adjusted slightly to compensate. This is effectively using fixed step size integral feedback control of the exposure. The feedback gain in adjusting the exposure must be sufficiently small to prevent instability resulting from natural differences in colour of the produce from item to item. The large dead band allowed for normal highlight pixel values (from 192 to 250) prevents changing the shutter speed at each image.

3.2 Segmenting Item from Background

At the same time as calculating the histogram, the contrast within the image is enhanced using a nonlinear lookup table. The purpose of this is to reduce the contrast range within the item of produce while enhancing the contrast with the background. The result is shown in figure 5.

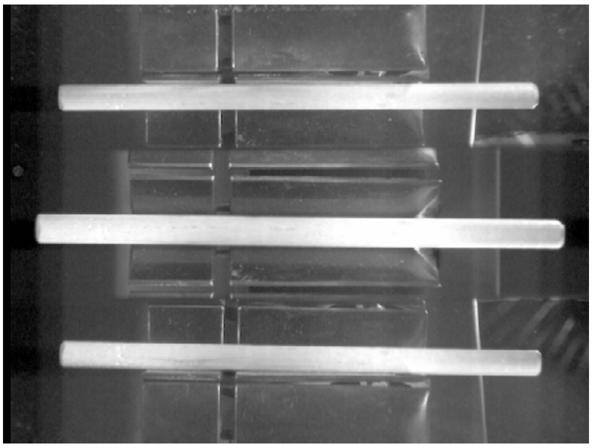

Figure 5: The image after contrast enhancement. The cups are clearly visible in the background, particularly the specular reflection from the right hand end.

The specular highlights may be removed by a greyscale closing operation [2] using a 1x9 pixel window. Rather than performing an erosion followed by a dilation, the closing may be efficiently performed in one pass by scanning to locate the local maxima and removing the peaks.

To remove much of the remaining background clutter, a background subtraction is performed. For this the image is split into three sections (one for each view). For each section the average is calculated for each column of pixels. The range of pixels in the column is then linearly expanded to set the background value to zero:

$$\hat{p}_i = \frac{p_i - \mu}{255 - \mu} \tag{6}$$

where μ is the column mean.

The final segmentation step is to threshold the image using a fixed threshold. There is some noise in the two side views as seen in figure 6.

A binary opening operation using a 9x1 element window is performed while thresholding to remove virtually all of the cup. Again, an efficient implementation is used – if there are

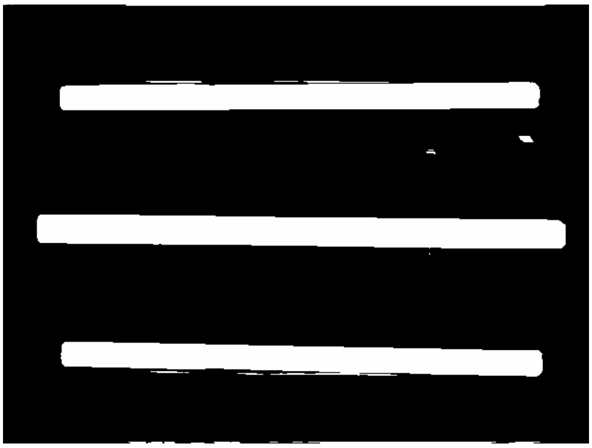

Figure 6: After thresholding, some cup remains in the side images. Much of this may be removed by filtering

fewer than 9 consecutive white pixels, in a column they are set to black. A similar opening and closing are performed using a horizontal 1x5 window to clean up the image.

3.3 Volume Measurement

At this point, the number of pixels in each column of the item can then be counted for each of the three views. From equation (2), a volume estimate can be calculated for each of the 3 views (although the central view will have a different scale factor to the side views).

With this approach, accuracy is limited to the diameter of the item in pixels. Since the edge of the produce is localised to the nearest pixel, at each position along the item, the measured diameter may be up to 1 pixel out. As a typical item of produce is only 30 to 40 pixels in diameter, this limits the accuracy to about 3%.

An improvement may be obtained by locating the edge of the produce to sub-pixel accuracy [3]. For this, a horizontal edge filter using a linear 3x3 kernel is applied to the image before thresholding. The results of this filter are shown in figure 7. The locations of the local minima and maxima can be estimated to a fraction of a pixel by fitting a parabola to the values of the extremum pixel and the values above and below. For efficiency, the thresholded image is used as a guide and the sub-pixel localisation is used to adjust the detected diameters.

While this will give an improvement of the estimate of the volume of each of the images, the error in estimating the weight from the volume will ultimately limit the accuracy obtained.

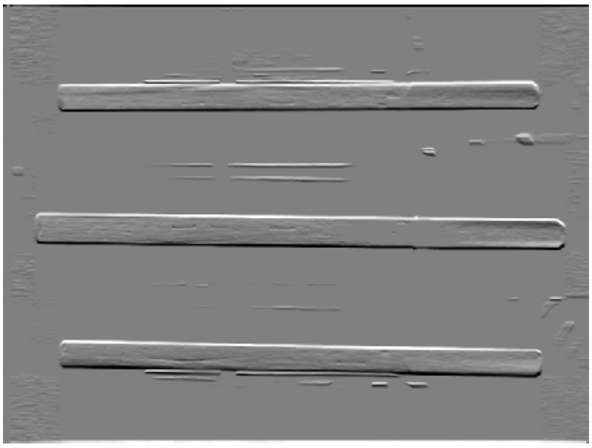

Figure 7: After edge detection. By fitting a parabola to the edge strength, the edge may be located to sub-pixel accuracy.

3.4 Calibration

There are two aspects requiring calibration. The first is the difference in scale between the central view and the two side views, and the second is the calculated volume (in pixel units) and the weight.

The scale factor between the centre and side views is estimated from the ratio of the length of the item in the centre view to the average length of the item in the two side views. This is averaged over a large number of items to give a robust value for the scale factor.

The item volume is calculated from the average of the volumes derived from each of the two side views. If estimating the volume was the end goal, a static calibration could be performed once the system was set up to determine the size of each pixel. However, for estimating the weight, such open loop calibration has a serious drawback. There are small, but consistent variations in product density throughout the season, and possibly throughout the day. To give a more accurate weight estimate, it is necessary to dynamically adjust the calibration as the density changes.

A closed loop calibration method can overcome such density variations as long as the change in density is gradual, and the density is remains consistent between individual items. Closed loop calibration is performed by weighing a random sample of items as they are processed. Every 100^{th} to 500^{th} item is directed down a chute with a load cell. This corresponds to 1 item every 10 to 30 seconds. This is sufficient time to obtain an accurate weight measurement from the load cell. The weights of the last 100 items are used to build a dynamic calibration curve through the measured points. This then provides the required mapping from the measured volume in pixel units to the estimated weight of each item. After being weighed, the calibration items are redirected back to the input of the system where they are reprocessed.

3.5 Quality Assessment

In addition to estimating the weight, the image analysis system is also used to perform a simple quality assessment of each item. The central view is more suited for grading based on any surface features present because the lighting can be arranged so that it is more uniform on the top surface. This is more difficult with the side views because the cup limits the light that can be brought onto the item. With the current arrangement, the visible surface only accounts for about 30% of the total surface of each item preventing complete surface inspection.

The produce may also be graded on the shape of the cross-section. Two views are insufficient to determine if the item is elliptical in cross section because if the axis of ellipticity is either parallel to, or perpendicular to the direct view, then the projections in each of the side views will have the same area. However, if the projected area in the direct view is significantly different from that of the off-axis views, then the object is elliptical. If the orientation of the elliptical cross-section is at other angles, then the range in measured areas gives an indication of deviation from circular cross-section. While there is little point in attempting to measure the actual ellipticity because the cross-section is seldom a true ellipse, a large range in measurements can be used to reject those items that deviate significantly from a circular cross-section.

4. IMPLEMENTATION

The control and vision system for this project have been implemented using LabVIEW. This has enabled different control and sorting algorithms to be simulated easily, and then used in the final production system without change. LabVIEW enabled the trigger, image capture and chute triggering to be easily integrated within a single system. It also incorporates the dynamic calibration by randomly selecting items to send to the load cell for weighing. The scales are also under the control of LabVIEW enabling the system to be completely integrated within a single environment.

LabVIEW also provides the tools for constructing a simple user interface, and for maintaining and providing statistics as the produce is graded. This is important as it enables the grower to be remunerated according to the quality of the produce provided, and provides a convenient mechanism for setting various parameters in the packing house when the system is set up.

All of the image processing routines are written as optimised C routines, and provided to LabVIEW using a DLL. These are then accessed through LabVIEW's DLL interface as custom operations. This approach was used rather than make use of the image processing operations provided by the LabVIEW image acquisition module because it enabled the operations to be specifically tailored and optimised for this application. The use of custom image analysis software has enabled several generic operations to be combined into a specific operation for this application. Algorithms for implementing standard operations have been modified to significantly reduce the processing time required.

Such optimisations have enabled the final implementation of the image analysis routines to process each image in 35 ms on a 1.8 MHz Pentium P4. This corresponds to a maximum processing rate of 28 items per second, which is well within the target rate of 15 items per second.

The accuracy of the weight estimate of each item is determined by a number of factors. These include variations in density of the object; errors in estimating the volume; and the accuracy of any calibration. The first two components each have an accuracy of approximately 2-3% as outlined earlier in the paper. The calibration dynamically associates the actual weight of each item with its measured volume, so is not expected to introduce any additional error. The total error in measuring any one item is therefore expected to be in the range 4-6%.

Table 1 summarises the results of weighing 2000 items. Overall, the RMS error in measuring each individual item was 4.4%. However, much of this error results from the smaller and underweight items. This results from the reduced number of pixels across the width of the item reducing the accuracy of the diameter estimation, and hence the volume.

However, when the bundle weights are considered, the central limit theorem will reduce the relative error. This is because errors of the individual items are independent. Since there are more of the lower weight items per bundle, this actually compensates for the greater errors in the individual items. Overall, the bundle error is about 2% in the final system.

5. SUMMARY AND CONCLUSION

More accurate grading is made possible by obtaining a better estimate of the weight of each item. Conventional weighing methods are very expensive for high speed grading and packing.

This paper demonstrates an approach to accurate, high speed weight estimation using image analysis. Two perpendicular views are utilised to obtain an estimate of the volume of each item, which is then related to the weight through a closed-loop calibration.

Table 1: Summary of results from weighing 2000 items. The items have been grouped here according to the target bundle size. The RMS Error is the error between the true measured weight and the weight estimated by the system. Error expresses the error as a percentage of the weight of individual items. The bundle error is the error accumulated over a 100 g bundle. This is lower than the item error because the errors within the individual items are independent.

Bundle size	Weight range	Count	Mean (g)	RMS Error	Error	Bundle error
Underweight	< 13.3 g	105	11.8 g	0.70 g	5.97%	2.05%
7 per bundle	13.3 – 15.4 g	169	14.4 g	0.65 g	4.52%	1.72%
6 per bundle	15.4 – 18.2 g	339	16.8 g	0.76 g	4.53%	1.86%
5 per bundle	18.2 – 22.2 g	493	20.3 g	0.83 g	4.10%	1.84%
4 per bundle	22.2 – 28.6 g	519	25.0 g	1.04 g	4.15%	2.07%
3 per bundle	28.6 – 40.0 g	326	32.6 g	1.26 g	3.85%	2.20%
Overweight	> 40.0 g	49	46.4 g	2.06 g	4.44%	3.02%

Processing speeds approaching 30 items per second have been demonstrated, with an overall accuracy of 4.4%. This corresponds to weighing individual items with an RMS error of 0.7 g to 2 g depending on the item weight. Accurate, high-speed weight measurement has enabled an improved chute allocation strategy that results in much tighter control over the weight of each bundle. Improved processing efficiencies have resulted in a higher packing throughput, and better control over the bundling process have resulted in significant cost savings for the company.

6. ACKNOWLEDGEMENTS

The authors would like to acknowledge funding for this project from Technology New Zealand through a Technology for Business Growth Grant.

7. REFERENCES

[1] Bailey, D.G., Hodgson, R.M., "VIPS - a digital image processing algorithm development environment", *Image and Vision Computing*, **6:** 176-184 (1988).

[2] Russ, J.C., "The Image Processing Handbook", CRC Press, Boca Raton, Florida (1995)

[3] Bailey, D.G., "Sub-pixel estimation of local extrema", *Proceedings of Image and Vision Computing New Zealand*, Palmerston North, pp 414-419 (November 2003)