

# Three-dimensional vision for real-time produce grading

Donald G Bailey<sup>1</sup>, Ken A Mercer<sup>1</sup>, Colin Plaw<sup>1</sup>, Ralph Ball<sup>2</sup>, Harvey Barraclough<sup>2</sup>

<sup>1</sup>Institute of Information Sciences and Technology

<sup>2</sup>Institute of Technology and Engineering

Massey University, Palmerston North, NEW ZEALAND

E-mail: [D.G.Bailey@massey.ac.nz](mailto:D.G.Bailey@massey.ac.nz)

## ABSTRACT

Produce is often sold by weight, so one of the roles of the grading system is to allocate each item to a particular chute for packing into fixed weight bundles. Accurate, high-speed weight measurement is difficult and expensive, so machine vision is used to estimate the weight of each item. Previous estimations relied on a single diameter measurement, which resulted in large errors. To ensure that the minimum weight was provided, each bundle was on average 30% overweight. By improving the accuracy of the estimation, and combining this with an improved chute allocation strategy, significant savings can be made.

The weight estimation in the system under development is based on the projected area of each item. The error in weight estimation was further improved by measuring the projected area from two perpendicular views. With the produce being sorted at a rate of 12 to 15 items per second, there are significant challenges in obtaining and processing the simultaneous perpendicular views of each item. The two views are captured of the item through the use of mirrors, and a third direct view is also obtained for quality grading purposes.

**Keywords:** 3-D imaging, projections, weight estimation, produce grading

## 1. INTRODUCTION

### 1.1. Produce grading and packing

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 of uniform size and quality. If the package weight is arbitrary, this is relatively easy to achieve, but for a fixed package weight 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 item is a small fraction of the total package weight, and either the items or the packages can be weighed while packing, this criterion can be achieved without giving away significant produce. 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. For the application of interest, the current grading speed is 8-10 items per second. It is desired to increase this to 12-15 items per second. These constraints make physically weighing the individual items extremely expensive, and impractical. To overcome these practical 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.

## **1.2. Existing systems**

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 unsuitable though if the produce is easily bruised because it relies on mechanical jostling to get the produce to fall through the mesh, and this may 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. Both of these constraints prevent a mesh based grading method in this application.

An alternative approach, using machine vision, is to make one or more measurements of the item being graded. 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 each item is approximately cylindrical. If only a part of the object is measured, it is assumed that the rest of the object is uniform, or can at least be predicted from the measurement made.

In this application, size, resolution, and speed constraints have meant that it was impractical to measure the whole object. The existing transport system consists of a conveyor with a cup or “V” shaped slot for each item. This ensures that items are presented individually to the imaging system for measurement. A disadvantage, however, is that it makes the segmentation of the entire object from the background very difficult. As each cup passes the camera, it triggers a high intensity, solid state flash unit to freeze the motion. The aperture of the camera is kept open and the next frame after the flash is triggered contains the required image.

Each item is approximately cylindrical, and the length is known (each item has been pre-trimmed to the same length) so a single diameter measurement is used to estimate the volume of the item. Making a single measurement simplifies the segmentation of the object from the conveyor system, because where the measurement is made, the item extends past the conveyor enabling the item to be imaged against a black background. The diameter is estimated by measuring the area of the object within the field of view (a fixed length along the object). Calibration then associates different diameter ranges with different weight classes of produce.

The system is calibrated through experience; with small adjustments made to the diameter thresholds until there are fewer than the allowed number of underweight bundles. The thresholds vary slightly during the season according to subtle changes in shape and density of the produce. The different “weight” classes have different numbers of items per package, varying from 3 to 7. Objects are graded by size – each package count has a diameter range that approximately corresponds to the desired object weight (from 1/3 to 1/7 of package weight).

## **1.3. Problems with existing system**

The biggest problem with the existing grading system is the very wide spread of weight in each package. The primary cause of this is that the whole item is not imaged, and the single diameter measurement is only a crude estimator of the weight of each item.

The aspect ratio of each item is such that to obtain an image across the length of each item, the resolution in measuring the diameter is significantly reduced. This is also coupled with the greater processing power required to segment the produce from the conveyor system. These constraints led to unavoidable compromises being made in the original system design.

A second limitation is that the weight of each item is not actually being estimated, but the produce separated into classes based on the diameter of each item. As a result, there is a significant spread in package weight. This is exacerbated by the low item count per package. With a large item count per package, the central limit theorem can significantly reduce the variance in package weight. However, when there are only a few items in each package, the central limit theorem has less of an effect because the probability of all of the items being at the low end (or high end) of the weight range is significant. This means that to avoid underselling, the packages must be significantly overweight on average. This combination of factors means that the existing system produces packages that are on average 20% to 30% overweight. As a consequence, 20% to 30% of the produce is effectively being given away.

The goal of this project is to significantly improve profitability by reducing the average package weight. Legal requirements mean that the minimum package weight must be maintained. (In practice, a small proportion of the packages may be underweight). This goal is achieved by improving the estimate of the weight of each item through more sophisticated image analysis techniques. The more accurate weight estimates allow more effective grading and package allocation strategies to be devised that significantly reduce the spread in package weight.

## **2. PROPOSED SYSTEM**

### **2.1. Mechanical system**

For practical reasons, it is necessary to keep as much of the mechanical system the same as the existing system. Therefore produce is still conveyed in cups past the vision system, where it is graded, and sent down an appropriate chute for packaging. Some changes are being made to the chute design to reduce the time between completing a package, and starting a new package. These changes are beyond the scope of this paper.

### **2.2. Weight estimation**

For the produce of interest, the density is reasonably consistent from item to item. Therefore, an accurate estimate of the volume will provide a good estimate of the item weight. From the volume of each item, the weight can be estimated to within 1% of the true weight.

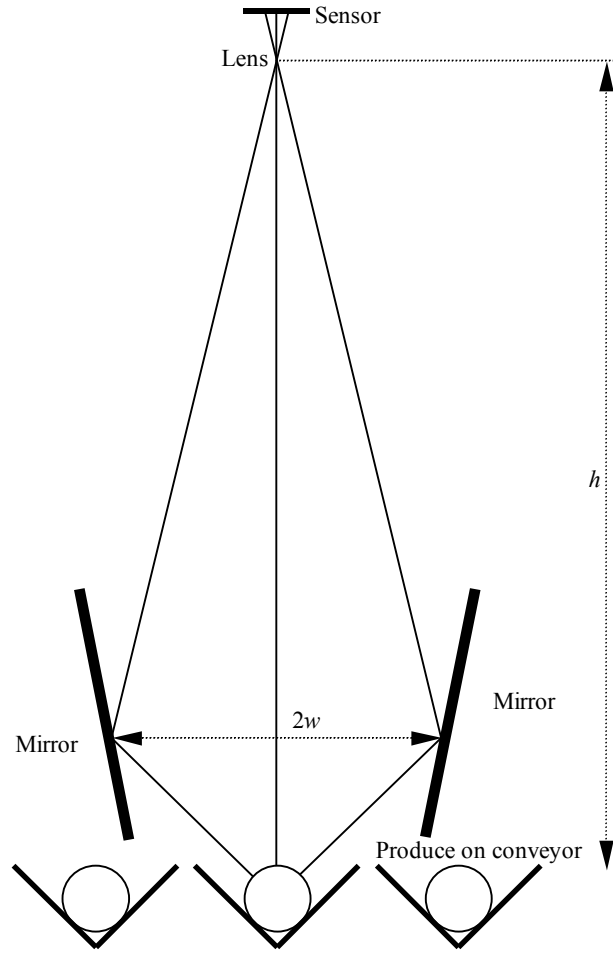
Although the produce being graded is approximately cylindrical, there is sufficient variation that a single projection does not enable the volume to be estimated with the required accuracy. Estimating the weight from one projection was only accurate to within about 10%. Therefore, while it will enable more accurate weight estimation than current methods, this is still not sufficiently accurate to exploit by improved chute allocation strategies.

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 tried was to image the end of each item, and assume that the cross section scaled along the length. 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. While this provides insufficient information to estimate the actual eccentricity, the volume could be predicted to within 2 to 3%. As this provided the best estimate of the weight of each item, this approach was considered for the prototype system. It also provides sufficient accuracy for a more sophisticated chute allocation strategy, as described in section 3.6 below.

The difficulty is then how to obtain two perpendicular views at right angles when the produce is moving along the conveyor at a rate of 12 to 15 items per second. The first alternative considered was to have two cameras mounted perpendicularly to the item at the trigger point. However, since the produce is significantly longer than its diameter, a single view will suffice, using a pair of mirrors to give the two views. This arrangement is illustrated in figure 1.

This setup not only gives two views at right angles, but also a third view directly. Although the path difference between the central view and the two off-axis views is not insignificant, if the camera is placed sufficiently far from the items (and a longer focal length lens is used) all three views will be in focus at the same time. (The effective path difference is slightly less than it appears because for the off axis views the sensor is further behind the lens).



**Figure 1:** Optical setup to obtain two projections at 90 degrees apart. A third projection is also obtained for free.

If the height of the camera above the direct view is  $h$ , and the space between the mirrors is  $2w$  then the off-axis path length,  $l$ , is given using Pythagoras's theorem.

$$l = \sqrt{(h-w)^2 + w^2} + \sqrt{2}w \quad (1)$$

The relative path length,  $R$ , scaled for the different angles using similar triangles is

$$R = \frac{l}{h} \cdot \frac{h-w}{\sqrt{(h-w)^2 + w^2}} \quad (2)$$

If  $k = w/h$  then

$$R = (1-k) \left( 1 + \frac{\sqrt{2}k}{\sqrt{1-2k+2k^2}} \right) \approx 1 + k \cdot (\sqrt{2} - 1) \quad (3)$$

To minimize the relative path difference, the mirrors should be as close together as possible (reducing  $k$ ). This means that the mirrors should be as low as possible to the produce. However the mirrors must also be sufficiently high that any

produce not positioned correctly within the cups on the conveyor will not foul the system. The relative path difference not only affects focus, but is also affects the relative scale between the off-axis images and the center image.

While in principle it may be possible to determine the ellipticity from the 3 projections, the formula is quite complex. In practice the exact formula is not particularly useful because the true cross-section is seldom an ellipse. However the variation in diameter between the three views does give an indication of deviation from cylindrical, and may be used as such for grading purposes.

### **2.3. Grading**

While the two side views provide information for size (and therefore weight) grading, having the third view is important for quality grading.

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 circular.

The central view is also 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. In this project, the produce is trimmed prior to grading, so the inspection at this stage is used to determine if sufficient stalk has been removed.

## **3. VISION TASKS**

### **3.1. 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. Lighting for the direct view is still relatively straight forward.

To achieve relatively uniform light distribution a pair of lights is used, aligned in the plane of the vertical axis, illuminating the item from the ends. If the lights are positioned above the item, some light is reflected under the item from the mirrors, improving the lighting in the off-axis views.

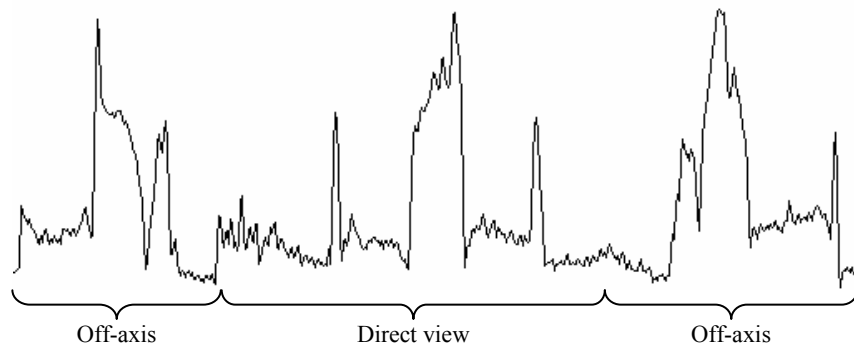
In the off-axis views, however, it can be more difficult to segment the bottom edge of the item from the cup. This is because the underside of the item is not directly illuminated. This makes it difficult to distinguish the item from the black of the cup. With the lights positioned above the item, the mirrors reflect some light onto the underside. Specular reflection from the edges of the cup actually make the black of the cup brighter than the object, and can actually be used to aid segmentation as described below.

### **3.2. Determining if an object is present**

When there is no item present in the cup, there is specular reflection from the bottom of the cup. This can make 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. If an item is present, the procedure continues with the rest of the processing.

### 3.3. Segmentation

Within the images captured, the cup and object are aligned horizontally within the frame. Each view occupies approximately one third of the image vertically, and the cup occupies one half of the image horizontally. If an item is present, it appears as light against a dark background, although the contrast on the side of the item against the side of the cup is in shadow for the off-axis views. This makes determining this edge of the item more difficult. Physical constraints make it very difficult to illuminate this area well. An intensity profile vertically through the image is shown in figure 2. This clearly shows the 3 views, and within each view, the item and the specular reflection from the edges of the cups.



**Figure 2:** Intensity profile across the 3 views.

Segmentation relies on the fact that although the cup is black, there is a specular reflection from the edge of the cup. The direct view is the simplest, with distinct edges well away from the edges of the cups. The distinct edges are detected using a linear edge detection filter. The two sides of the item consist of step edges of opposite polarity, with the region in between defined as the item of interest. The item should be approximately midway between the edges of the cup, as detected by the specular reflection peaks.

A similar approach is used for the off-axis views. These, however, are complicated by the fact that the underside of the item is adjacent to the specular reflection from the edge of the cup. The edge of the cup has a much stronger response to the edge detector than the edge of the object because of the shadowing. The cup edge can be removed by consideration by masking everything outside the distinct local minimum between the cup edge and the item. The top side of the item is better distinguished because it is better illuminated, and has a distinct edge well away from other strong reflections.

In applying the image processing steps to each item, the limited processing time is a significant constraint. Since image capture is triggered by the cup, there is only a small difference in cup location from one image to the next. However, after major adjustments, slight differences in the location of the camera, mirrors, and cup sensor may require a larger area to be searched. Within a run though, the consistency of triggering enables the cup edge locations to be predetermined during a calibration run, and appropriate masks constructed for removing the edges from the image.

The cup edges appear as lighter regions because of the specular component of the reflection. These are located by taking the central 50% of the image and averaging it horizontally, and then finding the local maxima in the vertical direction.

Given the masks of the edges of the cups, the region of the image that is processed can be significantly reduced. The edged detection starts at one end of the object, and when the initial edges are found, rather than process the whole image, the edge is tracked from one side of the image to the other. This local processing gives a significant speed improvement over processing the whole image each time.

### 3.4. Volume estimation

While each individual view provides an estimate of the size of the object, this is of limited accuracy because of potential ellipticity of each item. The average of the two areas provides a better estimator of the item size, and hence volume. The

areas within the off-axis views may be calculated as the edges are tracked. The average area is converted to a weight estimate using a calibration curve. This is described in more detail in the next section.

### 3.5. Calibration issues

Since each item is pretrimmed before imaging, the length of the items is constant. Therefore the measured projection area in pixels is proportional to the average diameter of the item imaged. The volume will therefore be proportional to the area squared. To estimate the weight from the projected area squared there are two constants. The first constant is to estimate the volume, and the second is an estimate of the item density, which is assumed to be constant within a batch.

Calibration is performed by weighing a random sample of items as they are processed. Starting with an approximate calibration curve, every 100<sup>th</sup> to 500<sup>th</sup> 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. This data is used to build a dynamic calibration curve through the measured points, and provides the required mapping from the measured areas 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.

If measurements are made on the direct view, a second calibration issue is that of the different scale between the direct and off-axis views. The central image is at a slightly higher resolution to the side images, as a result of camera geometry. If measurements made on the direct view are compared or combined with the side views, the difference in scale between the two, as given by equation (2) or (3) must be taken into consideration.

### 3.6. Size grading

Since the proposed system estimates the volume (hence the weight), it can allocate each item to a particular package rather than just a size category. For doing this, we have a target weight for each package as each item is added. For example for the 4 item in a package, if it is empty, it has a target weight of 25g. If it contains one item, its new target weight is 50g, if 2 items, the target is 75g, and if 3 items it requires only one more to complete the package, so the target weight is 100g.

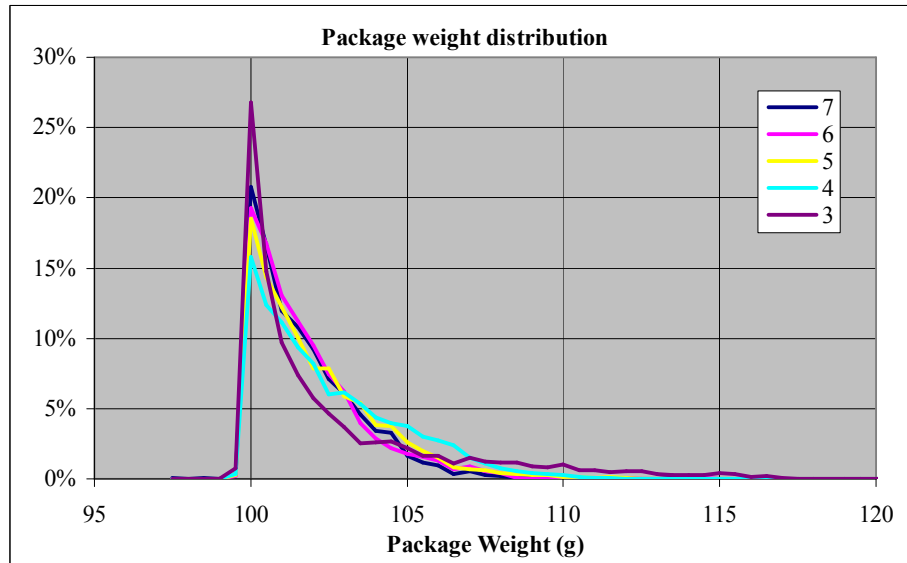
The item is allocated to the package that most closely achieves its target weight if the item was to be added to that package. When completing a package, if the final item makes the package underweight it is heavily penalized, to minimize light packages. The overall effect of this is to give a skewed distribution with very few underweight packages.

It is very difficult to get 3 item packages close to the target weight, since the ideal item weight is 33g. The next package size has a target item weight of 25g. As the count increases, the item weights become closer, and the items become more uniform, making it easier to achieve the target. For this reason, only 2 chutes are required for the 4 to 7 bundle sizes. The small count of the 3-item packages results in a variance that is significantly larger than the other counts. This is partially overcome by having 4 chutes for this size so that the individual items can be better matched to packages, reducing the overall variance. Figure 3 shows the results of simulations of this package allocation scheme given the assumption that the item weight is exact. In practice, the curves will be slightly more spread because of the errors in estimating the weight of each item. This will be more so for the packages with small item counts because the central limit theorem will have less of an effect in reducing the variance.

Underweight and overweight items are easily handled through the use of 2 item (for the oversize) and 10 item (for the undersize) packages.

**Table 1:** Package statistics of the proposed grading algorithm

Items per package	Average package weight	Proportion underweight	Proportion over 110 g
3	103.2 g	0.78 %	5.09 %
4	102.7 g	0.35 %	0.38 %
5	102.2 g	0.26 %	0.49 %
6	101.8 g	0.17 %	0.14 %
7	101.7 g	0.43 %	0.17 %



**Figure 3:** Distribution of package weights for different package sizes (simulated).

The results from figure 3 and table 1 clearly show the effect of the improved grading algorithm. The average package is less than 3.5% overweight, with very few underweight packages. The effect of small package size on the weight distribution is also evident. For small item counts, the tail becomes significantly longer, increasing the average weight. It should be noted, however, that these simulations assumed that the weight was known accurately. In practice, it is expected that the curves will become more spread, particularly for the smaller item count packages. This may require setting the target package weight to 102g or 103g to maintain an acceptable number of packages above the minimum weight.

#### 4. SOFTWARE IMPLEMENTATION

The control and vision system are being implemented using LabVIEW. This has enabled different control and sorting algorithms to be simulated easily. 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. All of the image processing routines are written as optimized C routines, and provided to LabVIEW using a DLL. These are then accessed through LabVIEW's DLL interface as custom operations.

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.

#### 5. SUMMARY AND CONCLUSION

More accurate grading is made possible by obtaining a better estimate of the weight of each item. This is achieved by using two perpendicular projections to obtain an accurate estimate of the volume of each item. Measurements made on a large number of items have shown that the produce of interest is uniform density within a batch, enabling the weight to be estimated to within 2% or 3%. This has enabled an improved chute allocation method aimed at minimizing the weight of each package. Each item is allocated to the chute that most closely achieves its target. As a result of this strategy, the average package is only 2% to 3% overweight given ideal weight estimation.

With the combination of these two, the average package is expected to be 5% overweight. This is a significant improvement over current grading methods, which give packages that are on average 20% to 30% overweight. As a result the improved grading method is expected to improve the profitability by 15 to 20%.