Artificial Intelligence Review **12:** 117–136, 1998.

Video Grading of Oranges in Real-Time

MICHAEL RECCE[1], ALESSIO PLEBE[2], JOHN TAYLOR[3] and GIUSEPPE TROPIANO[2]
[1]*Department of Computer and Information Science, New Jersey Institute of Technology, University Heights, Newark, New Jersey, USA (email:recce@cis.njit.edu)*
[2]*Agricultural Industrial Development (AID), Industrial Park, Blocco Palma I, Catania, Sicily (email: aid@cres.it)*
[3]*Department of Anatomy and Developmental Biology, University College London, Gower Street, London WC1E 6BT, U.K.*

Abstract. We describe a novel system for grading oranges into three quality bands, according to their surface characteristics. The system is designed to process fruit with a wide range of size (55–100 mm), shape (spherical to highly eccentric), surface coloration and defect markings. This application requires both high throughput (5–10 oranges per second) and complex pattern recognition. The grading is achieved by simultaneously imaging each item of fruit from six orthogonal directions as it is propelled through an inspection chamber. In order to achieve the required throughput, the system contains state-of-the-art processing hardware, a novel mechanical design, and three separate algorithmic components. One of the key improvements in this system is a method for recognising the point of stem attachment (the calyx) so that it can be distinguished from defects. A neural network classifier on rotation invariant transformations (Zernike moments) is used to recognise the radial colour variation that is shown to be a reliable signature of the stem region. The succession of oranges processed by the machine constitute a pipeline, so time saved in the processing of defect free oranges is used to provide additional time for other oranges. Initial results are presented from a performance analysis of this system.

Key words: pattern recognition, multi-processor, image processing

1. Introduction

Most of the processing of fresh fruit in packing houses is highly automated. Machines are used effectively for operations such as washing, waxing, sorting by size and color, and packing. However the most important step in the process, namely inspection and grading in quality, is still, with few exceptions, performed manually throughout the world.

This step has not yet been fully automated in production systems, as it requires fast and complex image analysis. Although grading of fruit shares several common features with more classical automated inspection of manufactured goods, this problem is significantly more difficult due to the wider range in variation found in natural products. Automation of the grading process is expected to reduce the cost of this important step and to lead

towards the standardization of grades of fruit that is desired by international markets. The inspection process must grade individual oranges into a small number of quality bands (typically three). Oranges are assessed according to surface characteristics including: discoloration, bruising and other blemishes. The grade is a measure of the number and size of these surface marks. The process is further complicated by the fact that, in an image, the stem is difficult to distinguish from defects. Part of the difficulty in correctly grading oranges is the process of detecting and distinguishing the stem from other potential marks that determine the appropriate grade.

Several manufacturers have developed machines for grading oranges, but none of these incorporates a stem detection mechanism, and in general the bin selection process is less than precise. So far none of the existing machines is able to match the requirements set by the market. We know of only one prior description of an orange grading machine in the literature (Maeda 1987), but similar image processing techniques have been used to grade a wide range of fruit and vegetables (see Tillett 1991 for a review).

This paper presents a new approach, largely based on the use of neural networks to achieve a more thorough analysis of the surface of oranges, including the detection of the stem. This analysis is aided by a new mechanical design which provides simultaneous images of the fruit from six orthogonal directions. More traditionally the grading decision is based on one or more views from a single direction as the inspected object is moved under a fixed camera (Davenel et al. 1988; Shearer and Payne 1990). One of the key aims addressed by this video grading machine is the real-time performance required by the application. The detailed video-based analysis is computationally demanding, and the problem is compounded by the high throughput required for a commercial grading machine.

The key to the solution is the use of a flexible process, in which time consuming computations are only invoked when necessary. In the first step of the process a fast global analysis is initially applied to every view of each orange. Further, more detailed and time-consuming analyses are issued only when required as a result of the previous stages. As a consequence the processing time for an orange is highly variable. In general, the process is faster for high quality oranges, and the time required scales roughly with fruit size. To take advantage of this time variation, the grading decision and activation of the separating device is delayed so that processing time can be best apportioned between the stored images of fruit currently in the pipeline. Since the largest percentage of fruit is of good quality, the time frame between adjacent oranges can be much smaller then the actual time spent to perform the most detailed analysis for a worst-case orange. The effect of this is to improve the throughput and overall grading performance (proportion of

oranges correctly graded). However, correct grading is not guaranteed in a case in which many consecutive oranges all exhibit difficult surface features.

Even with careful time allocation the target performance can only be achieved if the processors are sufficiently fast. The video grading machine incorporates a commercially available digital signal processor (DSP) and a new specialized neural network parallel processor currently under development at Philips Electronics Laboratories (Maudit et al. 1992).

All of the layers of analysis used by the grading algorithm are based on neural networks, fed by different feature extractors. This application is particularly well suited for neural network based techniques, because it is difficult to define any geometrical or spectral properties for the natural surface characteristics of fruit (stem, discoloration, bruising and other blemishes). However, the human eye is able to distinguish between a good orange, a defect, or a stem quite easily, which shows that the visual appearance, at modest resolution, contains the necessary information for classification.

2. System Overview

The automated fruit inspection machine described here consists of three principal parts: a system to convey the fruit into the vision chamber, the vision chamber itself, along with the image processing system, and finally the mechanism which diverts the oranges into different lanes according to their estimated quality (Figure 1). Although the most important part of the system is the image processing system, the method used to obtain the images is not at all trivial, and its development has been closely linked to the processing algorithms. The basic problem is to inspect, at high speed, the complete surface of a nearly spherical object. The wide range of sizes and shapes found in oranges preclude the use of conventional conveyors (rollers and belts). A novel design has been developed, permitting the simultaneous inspection of the entire surface of an orange during a short flight through the vision chamber.

2.1. *Impeller*

The first stage of this new system is a mechanism for accelerating the oranges to a suitable velocity, spacing them apart from each other, and propelling them along a fixed trajectory through the inspection chamber. The idea here is that each orange should follow roughly the same path independent of its geometrical properties. In the present design this occurs as long as the contribution of air resistance is minimal, the orange is not rotating when it

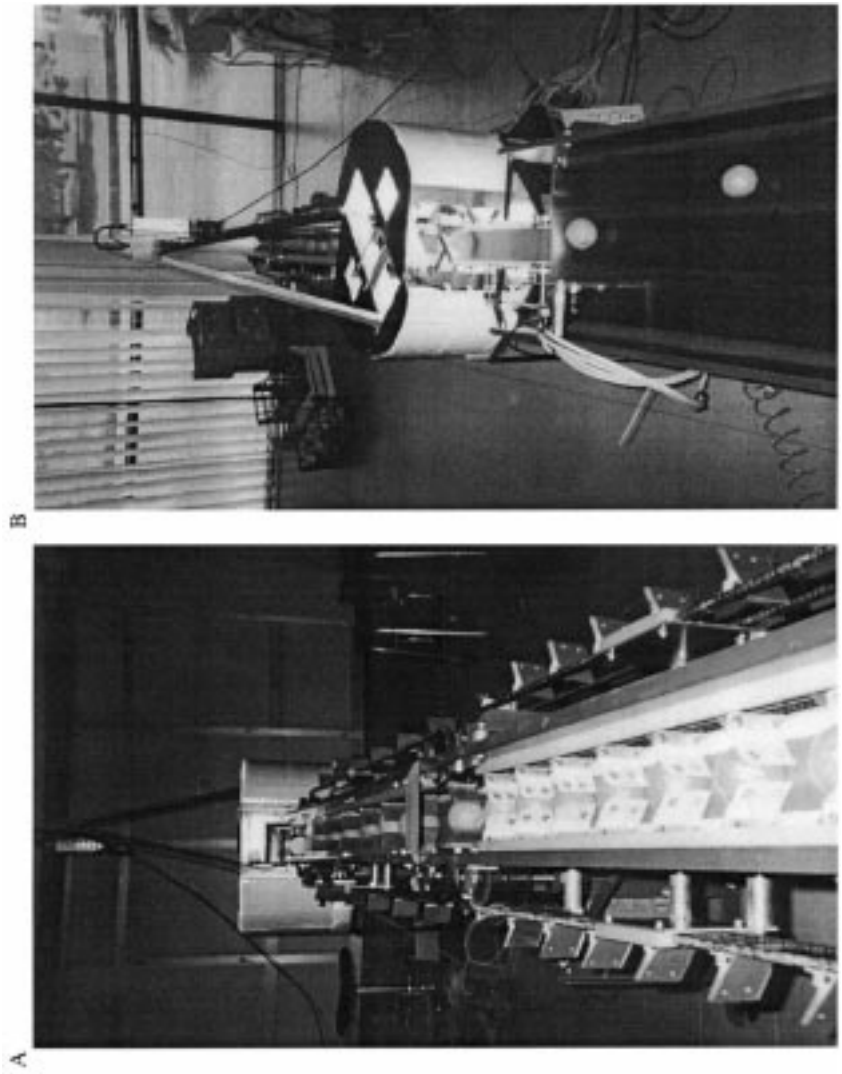

Figure 1. The prototype video grading machine. Part A shows oranges being carried by the impeller into the vision chamber. Note that there are two stages of acceleration, and in each of these the spacing between the oranges increases. Part B is a view of the oranges as the exit from the vision chamber and are sorted into three grades.

leaves the impeller and there is good control of all of the components of its initial velocity.

Also only one orange can be in the vision chamber at any point in time (otherwise it is impossible to obtain views of the front and the back of the orange). At prior stages in the packing house the oranges flow with a high packing density (approximately 100 mm center-to-center separation) at roughly ten oranges per second in an individual line. With a sequence of acceleration steps the speed of the oranges is increased by a factor of four and the spacing is increased by the same amount. The orange leaves the impeller with a speed of approximately 4 meters per second, travels 250 mm through the air, and lands on a belt that is moving with the same speed as the horizontal component of the initial velocity. The small change in velocity experienced by the orange minimizes the chances of bruising. The six views of each orange are captured simultaneously at the center of the vision chamber (see Figure 2).

2.2. *Image capture and processing*

The system hardware is VME-bus based, with four basic components: (1) the master processor (MC68030), (2) the color frame grabber based on a Texas Instruments TMS320C40 DSP, (3) the Philips prototype board, based on the L-Neuro2 parallel neural network processor and (4) an industrial digital interface board for synchronization and actuator commands.

The amount of processing required to make a decision on the quality of an orange depends on the quantity, type and distribution of defects. As described below a perfect orange can be processed very quickly, while a complex arrangement of defects requires a considerably longer time to process. For this reason we have constructed a multiple stage processing sequence, in which decisions that require the least time are performed first, effectively reducing the processing required at later stages. Furthermore, as described below the grading decision for an orange is delayed so that time saved in processing one orange can be applied to provide more time to process a subsequent orange.

There are three identified stages in the grading decision (identified below as separate processes). A fourth process provides the global supervision and control of the other three processes. The controlling process executes on the master processor (MC68030), supervises the other three processes, and keeps track of the location of individual oranges that are currently within the machine. The image processing steps are executed by the DSP and the L-Neuro2 (see (Maudit et al. 1992 for a description of the L-Neuro2).

In order to image the entire surface of the orange, one strategy is to use six planar views normal to the axes of a Cartesian coordinate system located at the

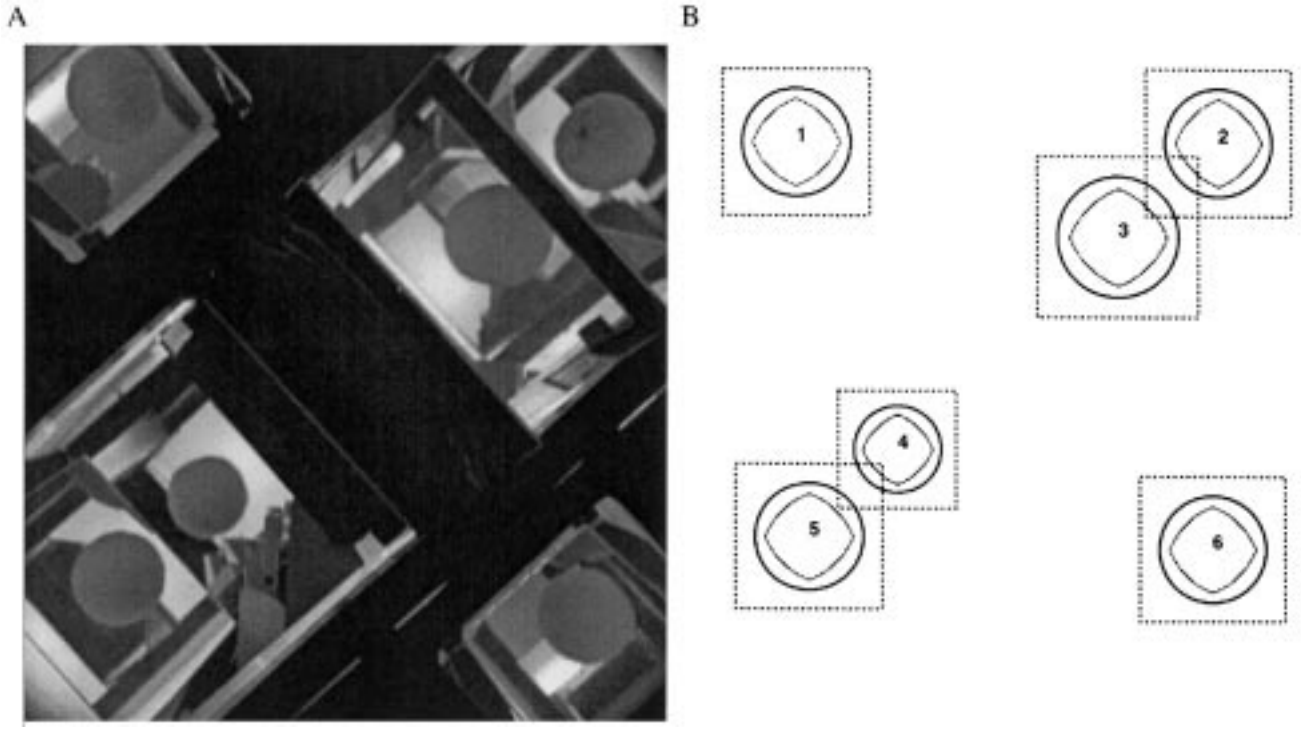

Figure 2. Six simultaneous views of an orange collected in the inspection chamber. The views are constructed using a set of mirrors. Part A shows an example image captured by the camera, and part B shows the regions extracted from the image. Color is used to segment the oranges from the background (Grasso and Recce 1996). In each view only the unique part is extracted as shown by the rounded shape that is drawn inside each of the circles.

center of the roughly spherical orange. The entire surface could be viewed, in theory, using two projections, but this results in considerable distortion from the curvature of the orange. If we assume that the orange is spherical than with six views the angle between the surface normal vector and the image plane is less than 45 degrees for over 95% of the surface of the orange. Six views are also quite practical to implement.

In the current prototype machine the six views are captured by projecting the image from a set of mirrors to an asynchronously triggered CCD camera. The camera is positioned directly above the center of the vision chamber, and the image capture is triggered when the orange is in mid flight in the center of the chamber. From the camera view point, four of the mirrors are in a "X" pattern and project four side views upwards. Two mirrors are used to project the underside of the orange upwards, and two mirrors are used to project the top view of the orange into the camera image. Figure 2 shows the camera view of the projected images. The image of the orange is the same size in each of the four side views, is slightly smaller (17%) in the underside view and is slightly larger (13%) in the top view. The figure also shows the region extracted from the six simultaneous views. The image chamber is uniformly illuminated with high speed fluorescent lights, but the chamber is designed so that light bulbs are not visible in any of the projected images.

In prior machines, a single image plane is used and the oranges are rolled and moved under the camera. In this case, the amount of surface seen, and the number of times that a surface element was seen, depended on the radius and eccentricity of the orange.

As soon as the image has been captured, the unique pixels are extracted from each of the six views of the orange. The symmetry of the viewing arrangement permits a calculation of the most appropriate regions from the six (assumed circular) views of the orange. The shape could be constructed by projecting the edges of a cube on to a circumscribed spherical screen, starting at a point at the center of the sphere, and then projecting this three dimensional curved cube onto a plane that is parallel to a face of the original cube. In addition, for each point selected from the image, a weighting factor is calculated which may be used to compensate for the viewing angle. Therefore, since points towards the edge of the orange are viewed more obliquely, a weighting value greater than 1.0 (ranging to approximately 1.6) is used to give each part of the orange surface equal value in calculations of the surface intensity histogram. The initial image extraction is performed by the TMS320C40 DSP.

Afterwards, each of the views is examined, and a view is excluded from further processing if it is unlikely that the particular view contains either defects or the stem of the orange. This decision is performed by a neural network algorithm that classifies color histograms (normalized red and green)

of the pixels contained in the view (described in more detail below). The DSP then passes each of the remaining views to the L-Neuro board for a more detailed analysis of local areas of the surface.

The image is scanned using a region based operator and each region is classified as defect free or not, using a second neural network (described below). The fraction of the surface area of the orange containing defects is the essential basis for the grading decision. Before treating an identified region as a defect, it is necessary to insure that it is not a stem. This is only done if the identified region is large enough to potentially be the stem. This is the most complex process in the grading operation and is carried out using a third neural network.

2.2.1. *Sorting system*

After the images of the orange have been obtained and the grading processing completed, the orange is deflected to the appropriate bin using purpose built pneumatic valves that release shaped jets of compressed air across the direction of motion and deflect the fruit into three separate areas.

During the visual analysis of each fruit, its dimensions, and hence mass are measured, which is used to set the timing and shape of the air jets. This activation is delayed to allow more fruit to be in a pipeline of process at the same time, and thus take advantage of the time-varying processing for each one.

3. Histogram Analysis

The first stage in the image processing is the classification of the color distribution of each of the views of an orange. Histograms of the normalized pixel values are constructed for two of the three color planes (red and green). Prior to this step, the color vectors are normalized to remove the effect of slight variations in the lighting. From empirical evidence, views of oranges that do not contain a stem or defects are found to have normally distributed red and green color components. This is to be expected, since the natural skin pigmentation is composed mostly of one color. Although, the peak color and its spread vary over different varieties of oranges, as well as over the season for the same variety.

Stems, discoloration, bruising and other blemishes tend to corrupt the smooth normal distribution of both the red and green pixel values. Unfortunately, this effect is often quite small, and the number of pixels with abnormal colors is unpredictable. Furthermore, in some cases a pigmentation gradient is present in good oranges, which may be caused by non-uniform exposure to sunlight during the ripening process.

Therefore, in part the classification of a view can be computed by the fit of the frequency distribution of the pixel values in the red and green color planes to a Gaussian function. Data which fit well to a Gaussian function are not likely to contain defects or the stem. The error is the summed difference between the best fit function and the histogram data. Figure 3 shows example best fit histograms for two defect free views, one view that contains a stem and one view that contains a defect. The examples in this figure show that while there is usually a good fit in the defect free images, there is a degree of variation. Also the view containing a stem and the view containing a defect can not be distinguished using the histogram analysis.

The surface defects on the oranges produce characteristic ranges of error in specific segments of the histogram. In order to evaluate the differences, the histogram data is divided into a set of segments, and in each segment I_i is computed to be the be the sum of the absolute value of the difference between the histogram data and the Gaussian fit to the data.

Empirically we found that a simple scoring rule based on the fit of a Gaussian function to the original histogram does not perform as well as a neural network based classifier that is used to learn the characteristic differences in a measured distribution from a normal distribution. The neural network algorithm used here is a modified form of back-propagation training of a multi-layer perceptron (Rumelhart et al. 1986; Werbos 1974) with ten inputs and two outputs. The input layer combines information from the red and green pixel color component histograms. There are two output classes: (1) no defects, or (2) either defects or a stem is present. Views of oranges that fall into the second class are passed to the second stage of the process, in which masks are used to analyze the surface in more detail. The number of input neurons, the partitioning of the histograms, and the number of hidden units have been empirically evaluated and compared.

As with other classification procedures, this procedure incorrectly classifies a fraction of the top quality oranges (Quality I) into a lower quality band (Quality II and Quality III), and places a fraction of the lower quality oranges into the top quality band. From the commercial point of view the full system should make fewer errors in downgrading the quality than in upgrading the quality. However, in this first stage of the processing it is safer to bias the process towards downgrading, and to run the additional processing stages when there is the possibility that a view of the orange contains a defect.

As suggested by Hush and Horne (1993), better generalization is achieved if the number of training samples is at least ten times larger then the number of weights in a multi-layer back propagation network. This neural network, and the neural networks at later stages of processing, is trained from a database that contains over 2000 examples (including a 20% test set).

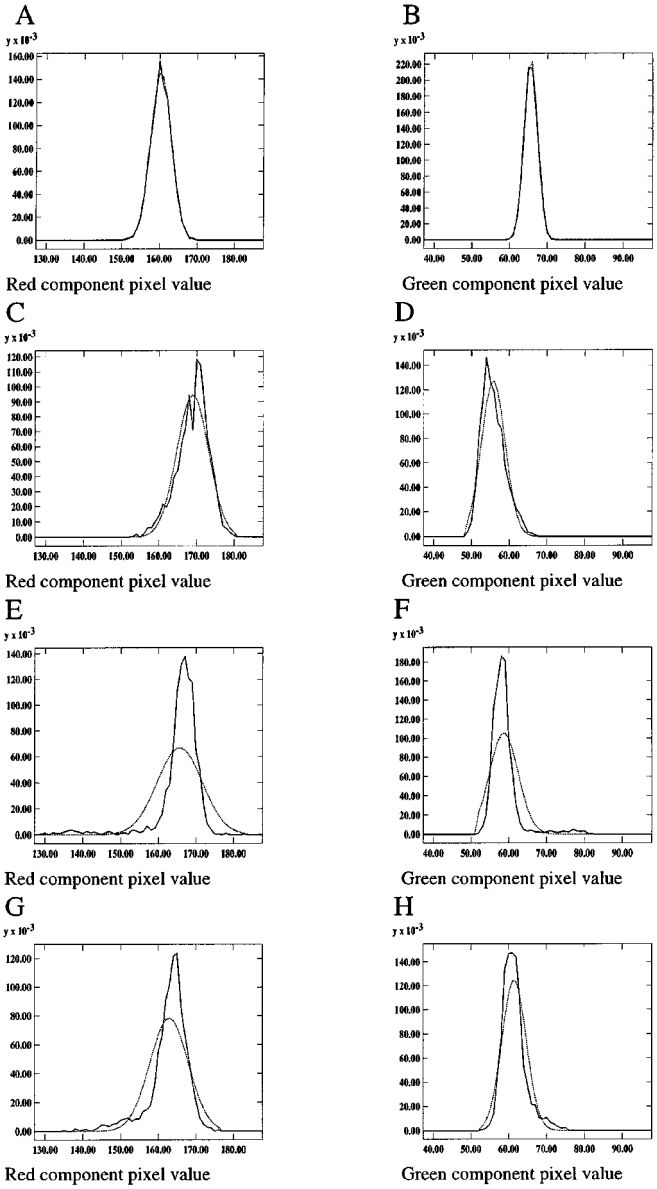

Figure 3. The red and green color distributions of four example views of oranges (solid lines), and the best fit Gaussian function (dotted line) to each of these distributions. Part A and B are the red and green pixel value distributions of a view that does not contain defects or the stem. Note that the distributions fit well to a Gaussian function. Part C and D contain the red and green color distributions for a second example view of an orange that does not contain defects or a stem. Note that the fit to a Gaussian function is less good, and that the mean and standard deviation of the two distributions are different for this example orange. Part E and F contain the color distributions and best fit Gaussian function for an image that contains a stem and part T and H are the color distributions for a view that contains a defect.

4. Local Defect Search

One approach to the identification of defects is to segment the defect regions from the parts of the surface that are defect free. A detailed analysis of the characteristics of defects leads directly to the observation that there is no regular clustering of defects in color space. Also in several cases, local color variation between pixels from the same orange is not sufficient to classify defects. The multiplicity of defect types has hampered prior attempts to segment the defects from unblemished surface regions, using local structural or textural properties. Furthermore, it is not clear that the computational effort required to identify all possible classes of defects, if it were possible, is justified, since all defect types contribute roughly equally to the final grading decision.

Instead we construct a simple and effective defect detector using a set of masks applied to regions of the image. The extracted local features are passed to a neural network based classifier that has been trained using a large database of samples. In almost all cases, a defect is characterized by a discontinuity in the skin pigmentation. But the color gradient can nevertheless be quite small, and similar to other non-defect sources of variation such as uneven illumination and border effects.

4.1. *Defect feature operators*

Ordinary edge detectors do not respond strongly to the appearance of surface defects on oranges. Instead, five larger and lower spatial frequency operators were used. In this process each image of the orange is divided into square regions that are the typical size of defects ($N \times N$), and all five operators are applied to both the red and the green color planes of these square regions. Each operator (M_k) is a $N \times N$ matrix, with integer elements that are either one, zero or minus one. The elements (m_{ij}^k) of the first four of these matrices are defined below:

$$m_{ij}^1 = \begin{cases} 1, & i < n/2 \\ -1, & otherwise \end{cases} \quad m_{ij}^2 = \begin{cases} 1, & j < n/2 \\ -1, & otherwise \end{cases}$$

$$m_{ij}^3 = \begin{cases} 1, & i > j \\ 0, & i = j \\ -1, & otherwise \end{cases} \quad m_{ij}^4 = \begin{cases} 1, & i < (N - j) \\ 0, & i = (N - j) \\ -1, & otherwise \end{cases}$$

In order to specify the elements of the fifth matrix (M_5) it is useful to define a set of square matrices, with size and with elements constructed in the same way as (M_k) except that $N' = \frac{1}{2}N$. M_5 is composed of four quadrants which are defined to be:

$$m_{ij}^{5} = \begin{pmatrix} -m_{ij}^{4'} & -m^{3'}ij \\ m_{ij}^{3'} & m_{ij}^{4'} \end{pmatrix}$$

The neural network based classifier is applied to each of the image regions. The input layer of the neural network contains ten neurons, computed by separately applying each of the five masks to the red and to the green components.

$$I_{k}^{r,g} = \sum_{i=1}^{n} \sum^{n} j = 1 x_{ij}^{r,g} m^{k} ij$$

The output layer of the neural network contains two neurons, and they are trained to classify defects, and used to estimate the severity of the defect.

The data base used to train this neural network based classifier is the same as the one described above. However in this case the parts of the image that correspond to defects or the stem have been extracted. Figure 4 contains part of the database of defects. Note that there are several examples of the same defect, which differ only in the selection of the start point for the sampled region. Instead of convolving the entire view with the defect detection masks, the computational overhead was reduced by applying the masks to tessellated regions of the image. In this case, the performance of the classification is significantly improved if the neural network algorithm is trained on images that have been translated in both directions.

Figure 5 contains samples of the stem data base. In this part of the analysis the stems are also classified as defects. This data base is then used in the next stage of processing to differentiate between the stems and defects. This stem detection process is only used if the stem has not already been identified in the prior processing of the six views of an orange.

5. Stem Detection

Stem detection is the most difficult step in the process of grading oranges, since the stem region looks quite similar to a defect. In general, global features such as the average color or the distribution histogram of red and green pixel color components, do not show significant differences between stems and defects. The strategy chosen here is to use a neural network based classifier, applied to identify spatial features extracted from the image region that contains the suspected stem.

The stem has a much more regular structure than the defects and has a high degree of radial symmetry, which can be used to aid in the identification

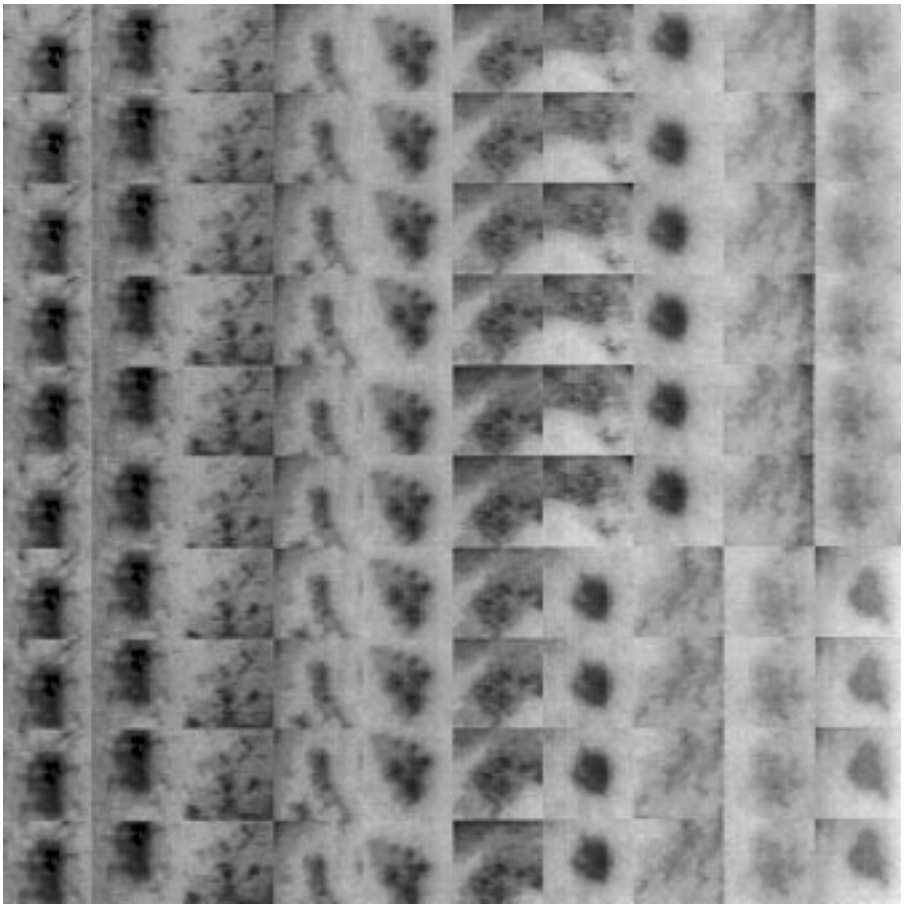

Figure 4. This figure contains 100 examples from the database that is used to train the mask based stem and defect detector. Note that individual defects are entered several times into the database with different displacements from the edge of the mask. The reason for this is explained in the text.

process. Often, but not always, there are characteristic radial groves that radiate from the stem attachment area. Also in some cases there are concentric rings around the stem attachment area. Unfortunately circular defects, of various types, also occur relatively often. We have found that the family of Zernike moments are very useful in detecting stems. In this process a set of Zernike moment masks are convolved with the region identified by the defect search.

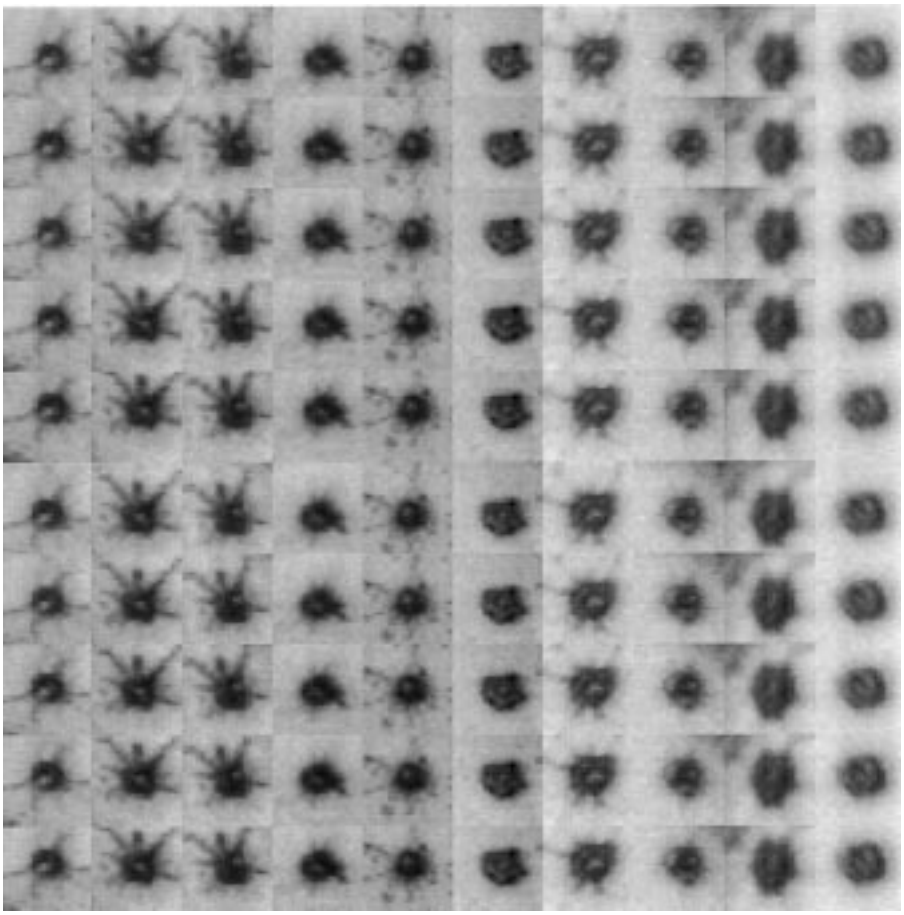

Figure 5. This figure contains 100 examples from the database of stems. This data base is used both to train the mask based defect and stem detection neural network, and is used to evaluate the performance of masks constructed from Zernike moments.

5.1. *Zernike moments*

When used as a convolution mask, the family of Zernike moments is particularly sensitive to circular symmetries. Importantly, Zernike moments are invariant under rotation, which makes this strategy computationally practical. Each of the views of an orange processed by this algorithm has already been scanned for defects using masks much smaller than the Zernike moment mask, so this stage is quite specific to suspected stem regions.

In these tests the radial size of the Zernike moments was not varied, since the size of stems of each particular variety of oranges is relatively fixed.

However, for a commercially viable machine, in which a large range of fruit sizes are processed, some scaling of the mask size may be necessary.

In order to generate a two dimensional square mask based on a Zernike polynomial, the mask is taken to have a unit radius in polar coordinates. Each pixel is assigned a complex value:

$$Z_{nm}(\theta, r) = [\cos(m\theta) + j * \sin(\theta)] * R_{nm}(r)$$

where: *n* is the major order of the Zernike polynomial, and *m* is the minor order. The radial variation of the Zernike polynomial results from the term, which is defined by:

$$R_{n,m} = \sum_{s=0}^{s=\frac{(n-m)}{2}} \frac{(-1)^s (n-s)!}{s!(\frac{n+m}{2}-2)!(\frac{n+m}{2}-2)!} r^{(n-2s)}$$

which has degree *n*, with only terms of the same parity as n. It generates an oscillating function with maximum wave number *n*/4. The minor order also affects the radial component, reducing the amplitude of the oscillations in the proximity of the circle center, while enlarging its wavelength. The border oscillations have larger amplitude and shorter wavelength as the value of the minor order approaches the value of the major order.

Figure 6 shows three examples of Zernike masks, represented in the picture with the intensity given by one component. More details on Zernike polynomials and their use as masks for image processing can be found in Khotanzad and Hong (1990).

In these masks, higher frequencies, both in radial and angular directions, vanish inside the mask. This occurs particularly near the center of the mask, where quantization effects are largest. Also the size of the window must be large enough to fully enclose the stem region. Reliable detection is only possible when the mask is well centered with respect to the stem region.

With the Zernike moments, the best input features for a neural network cannot be found empirically. For each Zernike polynomial major order *n* the possible minor orders are:

$$m = \begin{cases} n/2+1, & neven \\ (n+1)/2, & nodd \end{cases}$$

Even if the polynomial order is limited to be less than n_{max}, the number of candidate features (F) is still far too large for exhaustive search. Instead, it was decided to limit the evaluation to a search for the best set over a smaller number of candidate features. Four input neurons were used in this classifier, from 50 candidates required 200,300 runs against the database, which took approximately 100 hours on a SUN Sparc20 workstation.

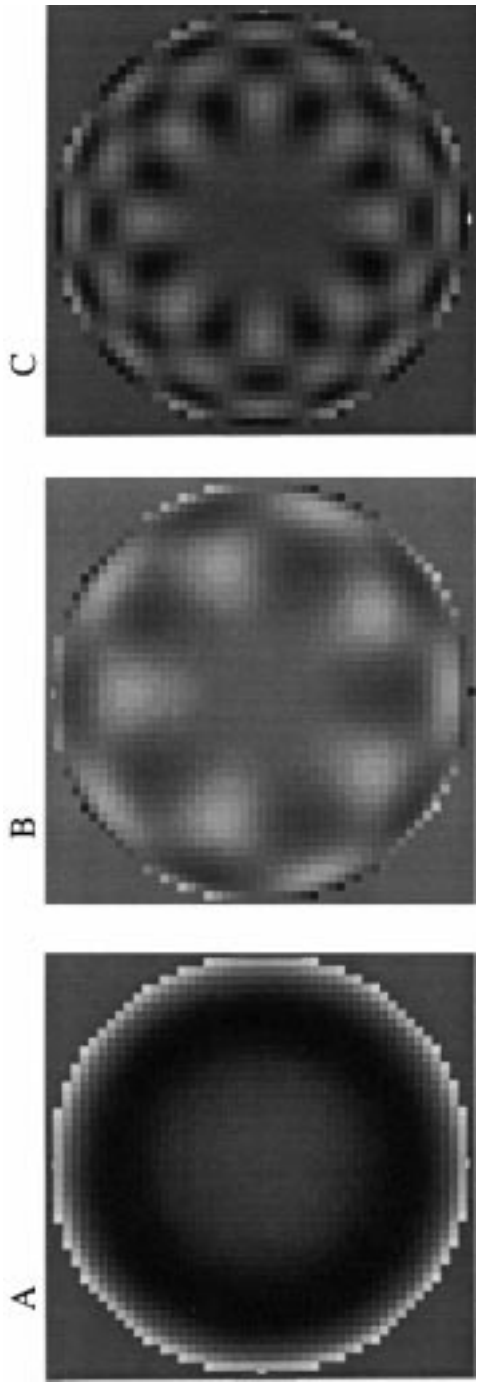

Figure 6. This figure contains three examples of Zernike moment masks. In each case the intensity corresponds to one component of the mask. Part A contains the modulus $Z_{5,3}$ (rank of 38.4 in the red color plane, 44.8 in green), part B contains the imaginary part of $Z_{9,5}$ (rank of 14.8 in the red color plane, 25.1 in the green) and part C contains the real part of $Z_{16,8}$ (rank of 53.3 in the red color plane and 16.8 in the green).

The screening of the subset of candidates was carried out by ranking all the features with a single-variate analysis. For each feature the rank (*i*) for a feature is computed as:

$$\frac{R_i = |m_{i,s} - m_{i,d}|}{s_{i,s} + s_{i,d}}$$

where

$$m = E[x]; s = E[(x - m)^2]$$

indexed over both the stems (*s*) and defects (*d*) in the database.

During the optimization process, ranking is done in the multivariate space. Classification during the analysis and test is accomplished by finding the features that have the smallest average distance to the masks. Figure 6 shows three example Zernike moment masks, and includes the rank of these masks for the red and the green color planes.

6. Results

The work described in this paper is directed towards developing a commercially viable machine with stringent real-time requirements. A successful machine should be cost-effective, robust, simple, and suited for implementation in parallel hardware. Nevertheless, the problem of inspecting fruit is intrinsically complex, and previous attempts to simplify it have been at best partially successful.

The flexible architecture of the algorithm presented here provides an advantage, by splitting the whole task into three simpler problems. The use of neural networks in solving all the steps of the grading has overcome some of the difficulties in defining a reliable description of the properties of the objects to be recognized and classified, which are intrinsically difficult to specify, and the results are so far within the expected performance of the video grading machine.

Figure 7 shows example results of the full grading process on four views of oranges. One of these views contains no defects, two of the views contain defects which have been detected, and the fourth view contains a detected stem.

The initial training and testing of the neural network algorithms was carried out off line using a database of images constructed by running oranges through the machine. The database contains images of 200 oranges of the tarocco variety, and includes 2296 images of stems and 2567 images of defects. Eighty percent of the database was used for training and twenty percent was

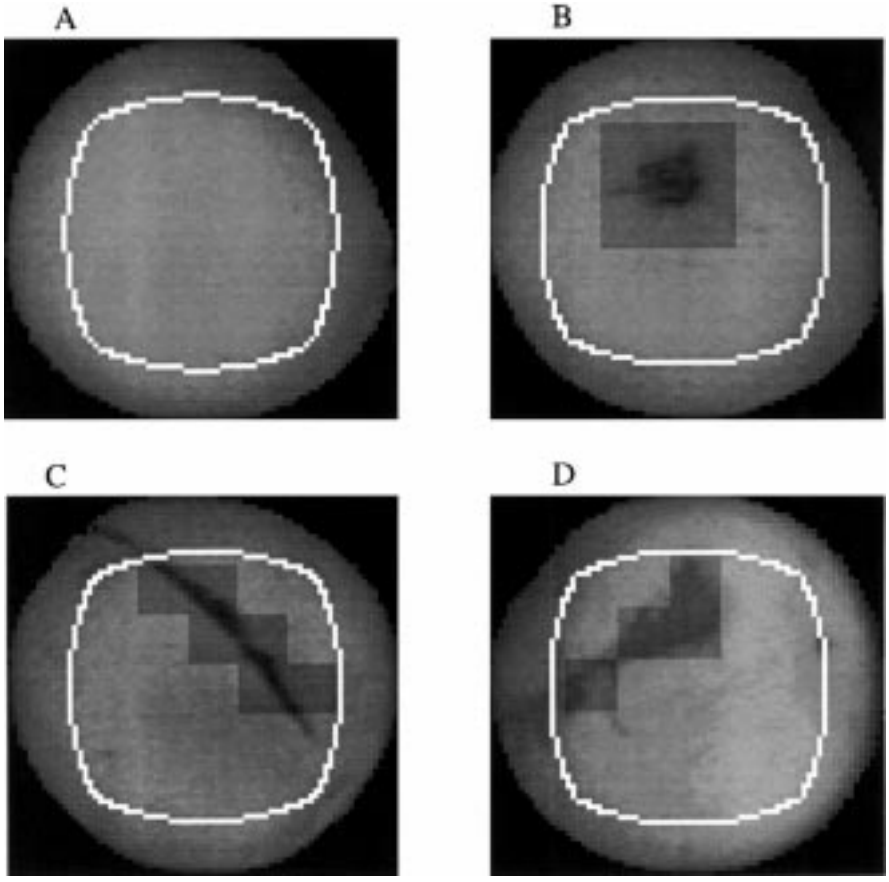

Figure 7. The combined result of applying the full set of processing stages on four example views. The first view (part A) does not contain defects or a stem. The second view (part B) contains a detected stem, and the remaining views (parts C and D) contain views with detected defects. The histogram for these three of the views (A,B,D) were presented in Figure 2 (A,B,E,F,G,H).

used to evaluate the performance of the algorithms. The error rates for each of the algorithmic steps are summarized in Table 1.

The errors in the columns labeled "$\rightarrow$ down" represent incorrect down grading of one of the views of an orange. For the histogram evaluation neural network, this means that further analysis is performed on an orange that is defect free. When the neural network used to detect defects erroneously down grades an orange than the region is searched for a stem. If the stem detection down grades the view then the stem is erroneously taken to be a defect. The errors listed in the table in the "$\rightarrow$ over" columns are false positives, in which defects and stems are passed undetected.

Table 1. The error rates of the three neural network algorithms. The " → down" colomns are errors that lead to down grading of views of oranges, and the "→ over" column are undetected defects

Network	Training set		Test set	
	→ down	→ over	→ down	→ over
Histogram	0.00%	0.00%	13.73%	2.06%
Defect	0.00%	0.62%	0.80%	2.02%
Stem	0.11%	0.87%	0.44%	1.31%

The imbalance in the two error classes is intentional, as described above. Moreover the differences in performance between the three neural network algorithms reflects the expected relative significance of errors in the final decision. It seems clear that the histogram analysis, which contains no spatial information, is unlikely to be refined significantly. Nevertheless, this is the processing step with the largest difference between the training and test errors, reflecting a rather poor level of generalization. This performance may improve with a larger number of oranges in the training set. For a given number of oranges in the database, the histogram training set is smaller than the data sets for the other two processing stages.

Evidence for the advantages provided by use of neural networks is shown here for the stem detection problem, which is the most important part of the whole algorithm. As a two class problem, the classical Fisher's linear discriminant seems a promising candidate technique (Duda and Hart 1973). When this technique is applied to the sample database, using the same reduced feature number used in the neural network results presented above, the results are quite poor. In this case 27.5% of the stems are not detected and 45.8% of the stems are incorrectly classified as defects.

This linear method may fail due to the implicit assumption that the distribution function of each feature is normal, and this is not true in the case of the Zernike convolutions over the database. In comparison, the neural network algorithm does not explicitly assume the characteristics of the probability distribution of the samples.

7. Summary

The present algorithm for fruit inspection, based on neural network classifiers, has come close to meeting the requirements of a future commercial video grading machine. Further work is envisaged to refine all stages of the classifier, and to balance the computational requirements of each of the stages. Also, thus far, little attempt has been made to use network architec-

tures other than back-propagation learning in a multi-layer perceptron. We plan to compare the performance or alternative non-linear classifiers. The present choice of approach has also been guided by the efficiency of running multi-layer perceptron algorithms on the currently available hardware. The machine has been tested with a few hundred oranges and a range of varieties. However, the machine may require a degree of profiling in order to cope with the large set of varieties of oranges that are commercially available. For example, the machine may need to be adjusted for particularly small or particularly large oranges. Also the machine may require the development of specialized sets of synaptic weights and other parameters in order to optimally grade particular varieties of oranges.

Acknowledgements

This work was partially supported by the European Community CAPRI Project, Sub-project PREFER, of the High Performance Computing and Network Program. The authors would like to thank Nick Ouroussoff and Dave Edwards for invaluable contributions to the design and construction of the video grading machine.

References

Davenel, A., Guizard, C.H., Labarre, T. & Sevila, F. (1988). Automatic Detection of Surface Defects on Fruit Using a Vision System. *J. Aric. Engng. Res.*: 1–9.

Duda, R. O. & Hart, P. E. (1973). *Pattern Classification and Scene Analysis*. Wiley-Interscience.

Grasso, G. & Recce, M. (1996). Scene Analysis for an Orange Picking Robot. In *Proc. of* 6^{th} *Inter. Congress for Computer Technology in Agriculture*, 275–280. VIAS.

Hush, D. R. & Horne, B. G. (1993). Progress in Supervised Neural Networks – What's New Since Lippmann? *IEEE Signal Processing*: 8–39.

Khotanzad, A. & Hong, Y. H. (1990). Invariant Image Recognition by Zernike Moments. *IEEE Trans. Pattern Analysis and Machine Intelligence*: 489–497.

Maeda, H. (1987). Grade and Colour Sorter Through Image Processing. In *Proc. of Intern. Symp. on Agric. Machinery and Intern. Coop. in High Tech. Era*, 350–356. University of Tokyo.

Maudit, N., Duranton, M., Gobert, J. & Sirat, J. A. (1992). Lneuro 1.0: A Piece of Hardware Lego for Building Neural Network Systems. *IEEE Trans. Neural Networks*: 414–422.

Rumelhart, D. E., Hinton, G. E. & Williams, R. J. (1986). Learning Internal Representations by Back-Propagating Errors. *Nature* **323**: 533–536.

Shearer, S. A. & Payne, F.A. (1988). Color and Defect Sorting of Bell Peppers Using a Machine Vision System. *Trans. ASAE*: 1–9.

Tillett, R.D. (1991). Image Analysis for Agricultural Processes: A Review of Potential Opportunities. *J. Aric. Engng Res.*: 247–258.

Werbos, P. J. (1974). *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences.* PhD thesis, Harvard University.