Lebensm.-Wiss. u.-Technol., 29, 203–209 (1996)

Machine Vision System for Automated Detection of Stained Pistachio Nuts

Tom Pearson

Western Regional Research Center, USDA-ARS, Albany CA 94710 (U.S.A.)

(Received May 1, 1995; accepted June 1, 1995)

A machine vision system was developed to separate stained pistachio nuts, which comprise about 5% of the California crop, from unstained nuts. Stained nuts have lower consumer acceptance and higher incidences of aflatoxin contamination. The machine vision system may be used as an automated quality control device to detect stained nuts. The system was tested on three different pistachio process streams: the bichromatic colour sorter reject stream (floaters and sinkers) and the manually sorted small nut shelling stock stream (floaters). The system had a minimum overall error rates of 13% for the sinker bichromatic reject stream, 14% for the floater bichromatic sorter reject stream and 15% for the floater small shelling stock stream.

Introduction

Most shell stains on pistachio nuts are caused by a hull defect known as an ‘early split’ (1). An early split is an abnormal opening on the suture of the hull while it is still growing on the tree, 3 to 6 weeks before harvest. Normally, the shell of a pistachio nut opens 2 or 3 weeks before harvest, and the hull will stay intact and serve as primary protection for the kernel. However, on about 2 to 4% of the nuts, the hull will split open with the shell. This opening leaves the kernel vulnerable to airborne mold spores and insect infestation. Early split nuts are more likely to be infested by insects and the aflatoxin-producing mold, *Aspergillus flavus*, than are normal nuts (2,3,4). Aflatoxin is a carcinogen that contaminates several other commodities, such as corn, cotton seed and peanuts. It has been estimated that the incidence of aflatoxin contamination greater than 200 ng/g in individual early split nuts is approximately 1 in 270, and the contamination incidence among all harvested nuts is between 1 in 21000 and 1 in 25000 nuts (4). It is expected that removal of stained pistachio nuts will remove most aflatoxin contaminated nuts, as well as nuts with an undesirable appearance.

Pistachio nuts are hulled almost immediately after delivery to the processing plant from the orchard. Next, they are immersed in a tank of water. Hull material, debris and approximately 10% of the pistachio nuts will float, while clean pistachio nuts generally sink. The lower density nuts that float usually have undeveloped kernels, low moisture content, hull adherence or insect infestation. Low moisture content and hull adherence to the shell indicate that a nut might have had an early split hull (2,5). The floating nuts are removed and processed separately from the sinking nuts. After the flotation tank, the nuts are dried down to about 6% moisture content and stored.

Before nuts are packaged, they are inspected by automated bichromatic ‘colour’ sorters to remove badly stained nuts. Rejected nuts are usually shelled, and the majority of the kernels are sold for use in prepared foods. Nuts accepted by the automated colour sorters are manually reinspected for final grading. Nuts with large amounts of staining, hull adherence or odd shapes are removed for shelling. Nuts that have moderate amounts of staining are dyed red or sold as a lower grade in-shell product. Nuts apparently infested with insects are discarded. Large, clean, unstained nuts are sold at a premium price.

Very little aflatoxin has been found in medium-to-large sized nuts that were accepted by both the bichromatic colour sorters and manual inspectors. Aflatoxin is primarily concentrated in four process streams: manually sorted dye stock floaters, bichromatic colour sorter reject floaters and small floater and sinker nuts weighing 0.57 to 0.71 g/nut (6). Approximately 50% of the nuts from these process streams have stains covering more than 10% of the shell surface area.

The bichromatic colour sorters determine the average colour of the nut. If a nut has a dark stain, the sorter will be likely to reject it. However, many acceptable nuts might have some minor stains near their stem, or they may have wide shell openings, exposing more of the dark kernel. These nuts will probably be rejected by the colour sorter. Early split nuts often have a long narrow stain adjacent to the shell suture, while the remainder of the shell may or may not be stained (1,5). These smaller, isolated stains might not cover enough shell area significantly to change the average colour of the nut shell. Thus, ‘suture-stained’ and other slightly

0023-6438/96/030203 + 07$18.00/0

stained nuts are likely to be accepted by the bichromatic colour sorters. Approximately 25 to 40% of all stained nuts are erroneously accepted by the bichromatic colour sorters. These errors result in extra labour required to hand-sort the nuts.

Instruments that have been developed to detect aflatoxin-contaminated pistachio nuts based on UV-fluorescence do not perform with adequate accuracy (7,8). There remains a need for a machine that has the resolution and computational power to separate nuts with small stains from unstained nuts with large kernel exposures. This will reduce the amount of manual labour involved in the final grading. Such a device might also be used to salvage uncontaminated nuts from some process streams that generally have high levels of aflatoxin contamination.

The objective was to develop an algorithm and machine vision system to separate stained nuts from nonstained nuts with a system inspection rate comparable to that of the current colour sorters in use (40 nuts/s). The cost of the system must also be comparable to current colour sorters (currently about $15000).

Methods and Materials

Data acquisition and processing system

To meet the inspection rate and cost constraints, a video inspection device designed for peanuts (ESM, Micro Scan II, Houston, TX) was modified (**Fig. 1**). This device uses three line-scan video cameras (256 pixels/line, 200 kHz pixel rate) to inspect the entire perimeter of the nut, as shown in **Fig. 2**. One pixel represents the intensity of an approximately 0.1 mm square on the nut. The image processing logic of the Micro Scan II approximates the area of a stain by counting the number of dark pixels. If the dark pixel count exceeds a certain threshold, the nut will be rejected. Key Technology (Walla Walla, WA), Delta Technology (Houston, TX) and other manufacturers also offer inspection systems, similar to the Micro Scan II, that perform thresholding and pixel-counting operations to recognize blemished products. However, because kernel exposure varies considerably on pistachio nuts, simple thresholding and pixel counting will not adequately identify stained nuts.

The Micro Scan II was modified by diverting the output of each camera to a digital signal processing (DSP) board (Dalanco Spry, model 310, Rochester, NY). This board is equipped with a 14 bit A/D converter (over a 10 volt range), a Texas Instruments TMS320C31 digital signal processor chip, and 32K of 32-bit word memory. The signal from the video cameras has a range of –2 and 0 volts. The video signal was not amplified because the 14 bit resolution from the A/D was considered more than adquate for the image processing.

The A/D converter is triggered by a signal from a photosensor that was supplied with the Micro Scan II. The photosensor activates a timer that triggers the A/D converter for 11.5 ms. This is sufficient time to capture fully seven frames of data from one nut. Most nuts travel out of view of the camera before six frames are taken. Only the very largest nuts (over 20 mm long) have data in all seven frames. The nut travels about 3 mm between each line-scan. The combined line-scan image represents roughly 4% of the entire surface area of the nut.

The DSP boards were mounted in vacant ISA slots in a 486, 33 MHz personal computer. Given the slow data exchange rates of this bus system, most of the image analysis is carried out on the DSP board between each pixel input. The TMS320C31 chip is capable of 16.5 million instructions per second (not including parallel operations). Pixel data is input at a rate of 200 kHz.

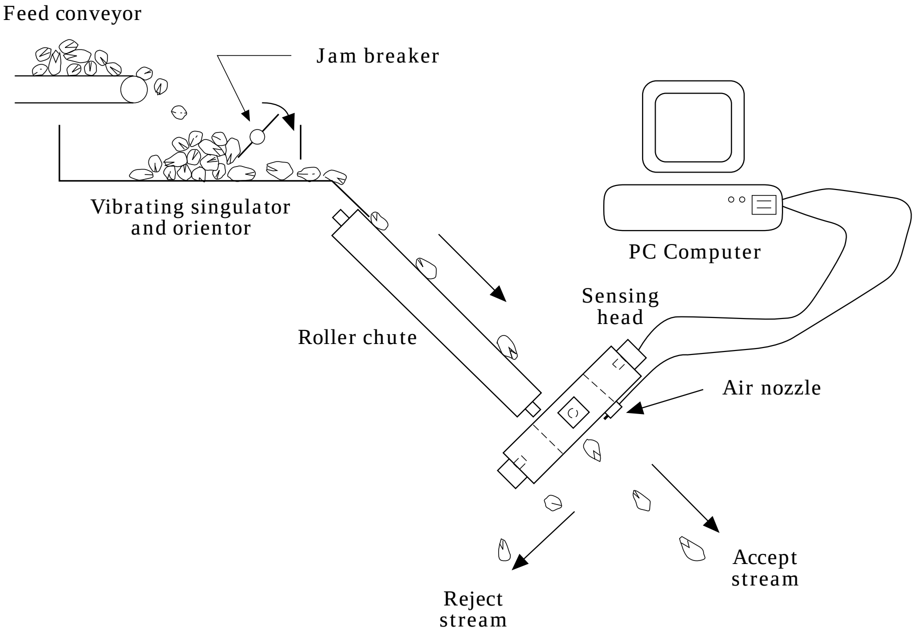

Fig. 1 Schematic diagram of machine vision system

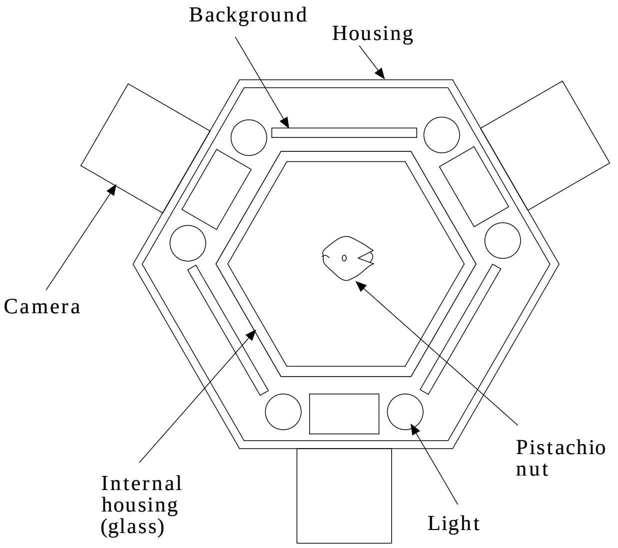

Fig. 2 Sensing head detail

Thus, 83 instructions by the DSP chip can be performed on the data between each pixel input. By performing image analysis between each pixel input, most of the image processing computations are complete the instant the nut travels out of view of the camera. This technique is necessary given the high product throughput required. The data from the three cameras are processed simultaneously by the independent DSP boards. After seven frames of video data from each camera have been processed, the results from each DSP board are transferred to the PC, where a final decision to accept or reject the nut is made (**Fig. 3**).

Camera background and filter selection

This system is intended to detect darker stains on the relatively light shell of the pistachio nut. In an ideal case, a band-pass filter in a wavelength range specific to the feature of interest's reflectance would be used. Thus, the camera response would be most sensitive for light wavelengths reflected from the feature of interest. Conversely, the background should be of a shade that reflects light wavelengths in a range similar to those reflected by an acceptable shell.

To measure the stain and clean shell reflectance characteristics, a spectrophotometer (NIR Systems, 6500, Silver Springs, MD) was used. Reflectance data in the visible light spectrum (400 to 700 nm) was measured in 4 nm increments on 30 stained and 30 nonstained nuts. Only an 8 mm diameter circle on the shell was exposed to the instrument. If a stained nut was being scanned, the nut was oriented to present the stain to the sensor. At the 658 to 662 nm band, the stained nuts had the largest difference in reflectance from the unstained nuts. This is probably due to the reddish colour of the stain. Thus, a filter with a band pass range from 630 to 690 nm was chosen.

Six different background shades were tested. Five of the backgrounds were shades of grey with visible light reflectances of 69, 74, 79, 84 and 89%, respectively. The sixth background was a latex paint (Glidden, shell white, Cleveland, OH) that matched the colour of unstained pistachio nut shells. With the band-pass filter installed on the camera, a background was mounted and 20 nuts were dropped by the camera. For each nut, the output signal from the camera was stored on a digital oscilloscope (Hitachi, VC-6155, Woodbury, NY). This was repeated for all of the six backgrounds to be tested. It was found that the largest difference in intensity between pixels representing the stained and unstained regions of the nut resulted with the pistachio-coloured background. This background was subsequently used.

Training set data acquisition

Samples of 63 stained and 63 unstained nuts were obtained from the bichromatic colour sorter reject floater stream of a California pistachio processor. These nuts were classified into stained and unstained groups by hand. A nut was considered stained if more than 10% of its shell area was discoloured. Individual nuts were fed through the system to acquire stored images for off-line analysis. The data acquisition system digitized the signals and stored them in the DSP board memory. After the nut passed the camera, the PC collected the data from the DSP boards and stored them in a separate file. The signal was converted from the A/D resolution of 14 bits to 8 bit resolution (0 to 255). The nut signal ranged from 190, which represented a clean shell, to 240, which represented the darkest stains.

Algorithm development

Figure 4 shows an example of one frame (256 pixels) of the video output signal from a stained nut and an unstained nut. Many frames from both stained and unstained nuts will show one peak that represents the dark kernel, which can be seen through the shell split. However, on stained nuts, the image line-scan may pass from a stained region to an unstained region several times. Thus, signals from stained nuts tend to have more discontinuities, or peaks and valleys, in their intensity profiles. Unstained nuts have large, light-coloured areas with very little variation. Thus, signals from nonstained nuts have more continuous, smooth, regions with lower intensities.

The change in intensity, or gradient, at each pixel was computed after performing a two-pixel mean filter on the image data. The gradient is directly proportional to the slope of the signal at a given pixel. The gradient at each pixel was computed by taking the difference in intensity between the two neighbouring pixels. The filtered image was only used for computing the gradients.

To quantify the characteristics of the camera signals, three parameters were computed. The first parameter indicates the number of major peaks, or high sloped regions (HSR). This was computed by counting the number of continuous regions of five pixels or more with a gradient absolute value greater than 1.50 and an intensity greater than 220. The intensity value of 220, gradient value of 1.50 and pixel count of 5 were chosen by trial and error. Several combinations of threshold values higher and lower than these were used to compute HSR for all of the training set images. For each set of threshold values used, the mean of HSR for stained and unstained nuts was subtracted. The final set of threshold values chosen produced the greatest difference in mean values.

The second and third parameters count individual pixels meeting certain criteria, rather than counting regions, as with HSR. The second parameter quantified the size of the low flat regions (LFR) in the signal. This was done by counting the pixels with a gradient of absolute value less than or equal to 0.5 and an intensity of less than 215. Finally, the third parameter quantified the overall number of discontinuities by counting pixels with a moderate slope (MS). To compute MS, pixels with a gradient absolute value greater than or equal to 2.0 and less than or equal to 3.5 were counted. The final set of threshold values to compute LFR and MS were determined by trial and error similar to the procedure used with HSR, by choosing the set that resulted in the greatest difference in means between the stained and unstained nuts.

Discriminate analysis (9) was performed on all possible models using LFR, HSR and MS. The full model,

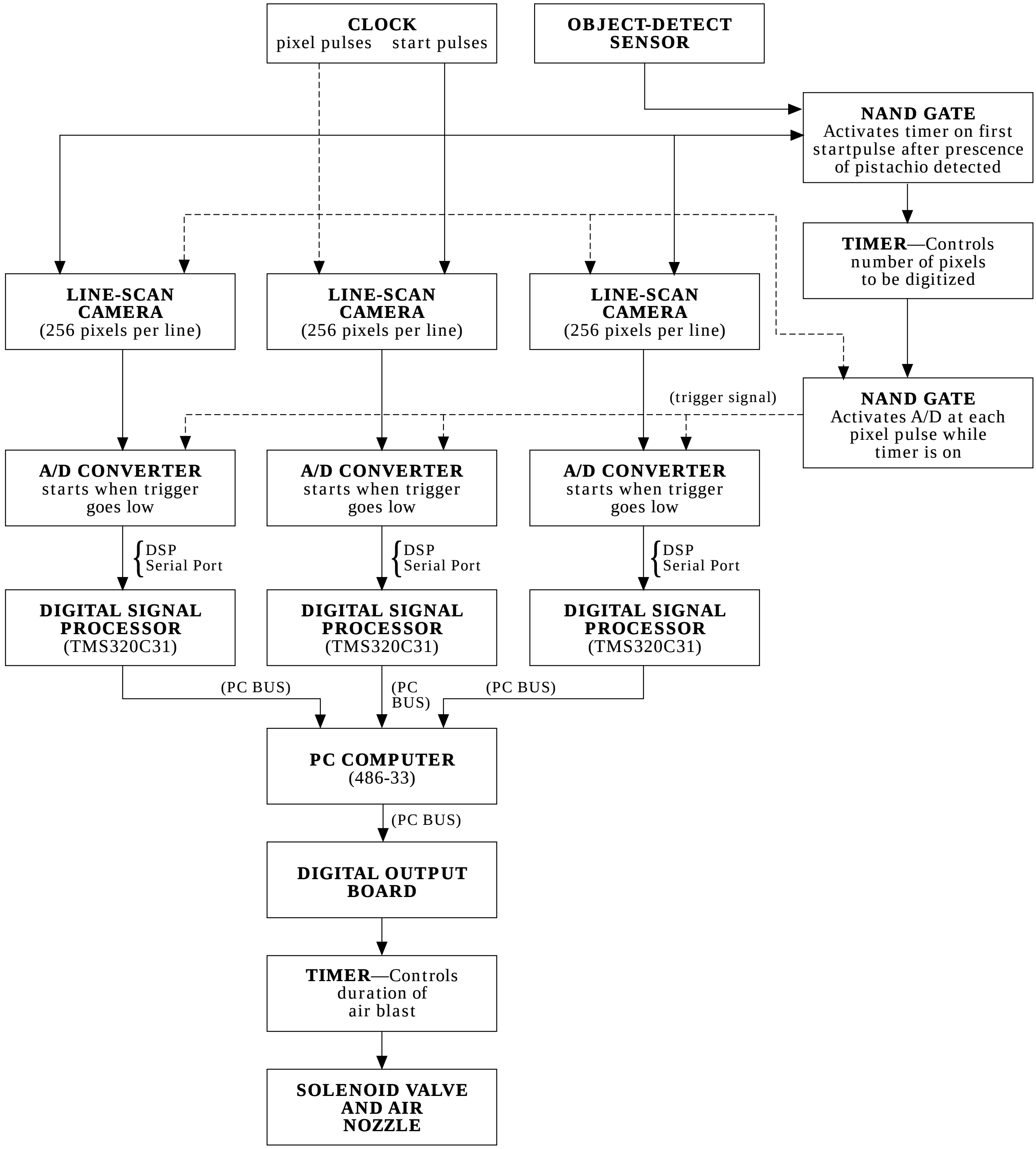

Fig. 3 Flow chart of machine vision system

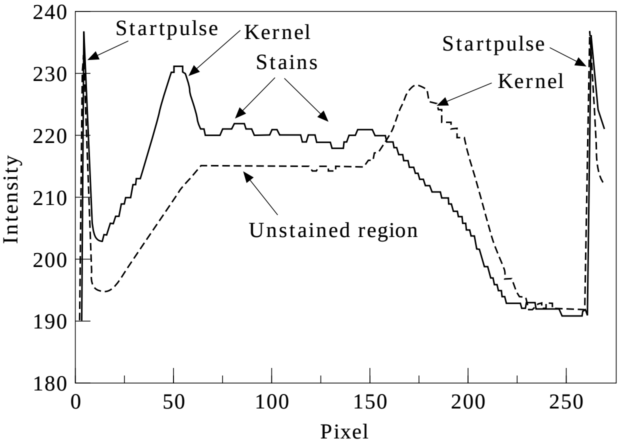

Fig. 4 Example of one line-scan from stained and unstained nuts. (-----) Unstained nuts; (———) Stained nuts

consisting of all three variables, had the highest classification accuracy. The within-group co-variance matrices with the full model were determined to be homogenous by the SAS discriminate analysis procedure. This determination then yields a linear discriminate function, as opposed to a quadratic function if the co-variance matrices were not homogenous. This is beneficial in that the linear function requires less time to compute. The model consisting of LFR, MS and HSR was used for further implementation of the algorithm. However, as will be discussed later, the actual discriminate functions implemented were developed from a larger sample set.

The algorithm to compute HSR, MS and LFR was coded in TMS320 native assembly code for implementation. The program requires up to 75 DSP chip instructions to perform the necessary computations between each pixel input. The PC polls a memory location on a DSP board to determine whether the processing is complete. After seven line-scans (1792 pixels) of data are processed by the DSP, the PC is notified, via the polled memory location, that new data are ready. The values of LFR, HSR and MS from each camera are transferred to the PC, and the corresponding data from each camera are combined.

Using the combined LFR, HSR and MS data, separate discriminate functions were computed for two floating nut process streams: the bichromatic colour sorter reject stream and the small shelling stock processing stream. Schatzki *et al.* (6) found the bichromatic colour sorter reject floaters to be heavily contaminated with aflatoxin. The small shelling stock nuts were accepted by the bichromatic colour sorter and classified by the manual sorters. The bichromatic colour sorter reject nuts had an average mass of 1.19 g/nut and the small shelling stock nuts had an average mass of 0.87 g/nut.

A third set of discriminate functions was computed for sinking nuts rejected by the bichromatic colour sorter (weighing 1.26 g/nut). These nuts were classified in three ways: stained, moderately stained and unstained. Moderately stained nuts are a subset of the stained nut group, they comprise nuts with stains covering 10 to 50% of the shell area. The classification is performed by passing nuts through the system twice. The first pass removes the stained nuts. The accepted nuts are inspected again (second pass) to separate the moderately stained from the unstained nuts. Thus two sets of discriminate function are required. The first set classifies stained nuts from moderately stained and unstained nuts, the second distinguishes moderately stained from unstained nuts. A data set containing samples of 300 stained, 300 unstained nuts and 300 moderately stained (bichromatic reject sinkers only) nuts were used from each processing stream. Nuts were classified as stained or unstained by hand, following the same protocol used with the training set. Once the discriminate functions were determined, they were programmed into the PC to automatically classify each nut.

Feeding system

The vibratory feeding system supplied with the Micro Scan II sorter was designed to handle peanuts and did not work properly with pistachio nuts. Many of the nuts would tumble down the rollers due to this system. Thus, an attachment to the vibrator was constructed to deliver the nuts to the rollers and cameras in a consistent orientation. The device singulates the nuts from a bulk delivery system by use of a vibrating hopper equipped with a rotating paddle to break jams at the output. The nuts are then conveyed to the rollers in a vibrating V trough (see **Fig. 1**). During this transport, the nuts orient themselves so their long axis is parallel with their direction of travel. In this orientation, nuts slide down to the rollers for delivery to the cameras.

The angle of inclination of the rollers must be steep enough to maintain the 40 nut/s throughput rate and create a gap between the nuts so that they can be properly ejected, yet it is desirable to minimize the speed of the nuts through the sensing head so they will be exposed to the cameras longer and to simplify the rejection device. This requires the slope of the rollers to be minimized. The roller inclination was adjusted from 30 to 60 ° above the horizontal. It was found, by trial and error, that 43 ° was the optimal angle of inclination. If the rollers are inclined more steeply, the nuts are more likely to tumble. At lower levels of inclination, the nuts sometimes touch as they enter the sensing head.

To measure the speed of nuts through the sensing head, a photosensor (Archer, 276-1657, Fort Worth, TX) was placed 50 mm downstream and upstream from the camera centre line. The time interval between detection of each nut by each photosensor was recorded with a digital oscilloscope (Hitachi, VC-6155, Woodbury, NY). The average speed of the nut was then computed. This was performed for 30 stained and 30 unstained nuts.

Air reject system

When a nut is determined to be stained, the PC computer outputs a signal through a digital I/O card (Keithly Metrabyte, PIO-24). This signal activates a timer that controls the duration of the air blast. The timer activates a power transistor to energize the air nozzle. To test this portion of the system, 100 stained and 100 unstained nuts were fed through the machine and inspected. The DSP and PC performed all the necessary computations to classify the nut. However, the PC was programmed to reject every nut that passed by the cameras, regardless of its stain type. The percentage of nuts successfully rejected was then computed.

Results and Discussion

The discriminate functions are used to compute the squared distance from an observation to the stained, unstained or moderately stained groups. The probability of a nut being stained is then computed using Eqn [1]:

$$\Pr\,(\text{stain}) = \frac{\exp\left(-\tfrac{1}{2}D^2_{\text{st}}\right)}{\exp\left(-\tfrac{1}{2}D^2_{\text{st}}\right) + \exp\left(-\tfrac{1}{2}D^2_{\text{un}}\right)} \qquad \text{Eqn [1]}$$

where $D^2{}_{\text{st}}$ and $D^2{}_{\text{un}}$ are the squared distances from an observation to the centre of the stained and unstained groups, respectively. If moderately stained nuts were being sorted from unstained nuts, the distance from the stained group, D_{st}, would be replaced by the distance to the unstained group, D_{md}.

A higher result computed from Eqn [1] represents a higher probability that a nut is stained. To classify nuts, a probability threshold level is chosen. Nuts with Pr(stain) values greater than the threshold value are classified as stained. The classification performance at any probability threshold is displayed in **Figs 5** and **6** for the bichromatic colour sorter reject stream and the small shelling stock stream, respectively.

The classification error rate for any number of groups, n, is computed by Eqn [2]:

$$\text{error rate} = \sum_{\gamma=1}^{n} \varepsilon_\gamma p_\gamma \qquad \text{Eqn [2]}$$

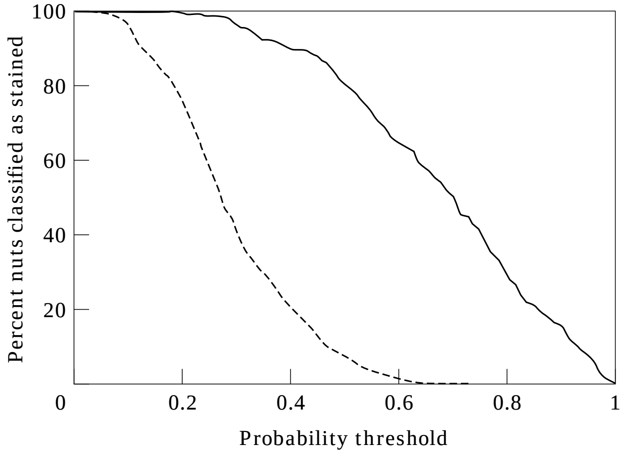

Fig. 5 Classification performance for bichromatic colour sorter reject floater nuts. (-----) Unstained nuts; (———) Stained nuts

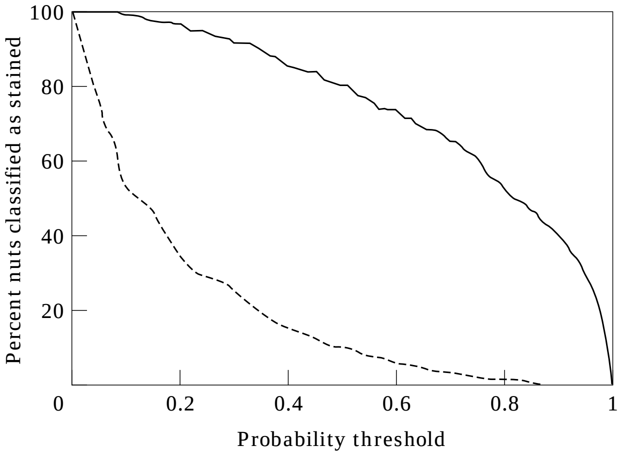

Fig. 6 Classification performance for small shelling stock floater nuts. (———) Stained nuts; (-----) Unstained nuts

where ε_γ is the fraction of group γ samples that were not classified into group γ, and $p\gamma$ is the fraction of the total number of samples (from all groups) that belong to group γ. For all of the process streams tested, the number of stained nuts approximately equalled the number of unstained nuts. The minimum classification error rate always occurs when nuts are classified using a 0.50 probability threshold. As can be seen from the classification performance curve for floater bichromatic sorter reject nuts (see **Fig. 5**) 20% of the stained nuts and 8% of the unstained nuts are incorrectly classified at the 0.50 threshold level. Thus, the minimum error rate for the floater bichromatic reject nuts is 14% ($[0.20 \times 0.50] + [0.08 \times 0.50] = 0.14$). The minimum error rate for small shelling stock sinkers is found in a similar manner and is 15%.

For the two-step sorting operation performed on the sinker bichromatic colour sorter reject nuts, the overall error rate was 13% on the 900 nuts used in the discriminate analysis. A different sample of 1000 randomly selected nuts from this process stream was sorted by the system and later manually inspected for classification accuracy. Of the nuts classified as stained by the machine vision system, 3% were observed to belong to the moderately stained group. Of the moderately stained nuts, 6% were found to be stained and 10% were found to be unstained. Finally, of the unstained nuts, 15% were considered to belong to the moderately stained group. The machine vision system classified 35% of the nuts as stained, 32% as moderately stained and 33% as unstained. Using these percentages and observed classification errors for the three groups, the system had an overall observed error rate of 11%.

The moderately stained and unstained nuts appear to be usable as in-shell product, although many of these nuts would require dying. These results indicate that the machine vision system can inspect the sinker bichromatic colour sorter reject nuts, which are destined for shelling, and salvage approximately 65% of them for sale as in-shell product.

Based on the stains, hull adherence and nut size in the stained nut groups of all three processing streams tested, many of the stained nuts would probably have had early split hull defects or insect infestation. These observations support the hypothesis that this machine vision system will separate nuts contaminated with aflatoxin from uncontaminated nuts. However, since the incidence of contamination is so low, Schatzki *et al.* (6) estimate that the quantity of nuts required to prove this hypothesis would be prohibitive. Approximately 1000 kg would be needed to be sorted and analysed for aflatoxin to draw statistically significant conclusions on the aflatoxin levels of a group of nuts.

For all three of the process streams tested, nearly all of the nuts with hull adherence were detected. The coarse texture of the hull causes high HSR and MS counts. This result indicates that this system may be useful to automatically separate dye stock from shelling stock. The algorithm is most sensitive to variations in intensity. As a result, false-negatives, or stained nuts that are not rejected, usually have only one stained region of a consistent intensity.

The nuts were found to be travelling at speeds ranging from 1.67 to 1.91 m/s, with an average speed of 1.78 m/s. There was no appreciable difference in speed between stained and unstained nuts. However, the inconsistent speeds of the nuts make stain area or nut size computations more difficult and probably less reliable. There is a 13% variation in the range of speeds measured; this could also lead to a 13% error in size estimation from the image data. Probably more accurate estimates on LFR, HSR and MS could be achieved if the nut speed were measured as it passed by the cameras. This may improve the overall classification accuracy.

The air-reject system performs at a very low error rate. When the PC was set to reject nuts regardless of their stain type, 100% of the nuts were successfully rejected at nut feed-rates up to 40 nuts/s. The performance deteriorated above the 40 nuts/s rate, and at 50 nuts/s, approximately 25% of the nuts were not rejected. The DSP and PC performed all the computations needed to determine a nut's stain type during this test. Therefore, no difference in performance should arise when only the stained nuts are to be rejected.

Conclusion

A working machine vision system has been developed that separates stained from unstained nuts. The machine may be used to reduce the amount of manual labour in grading pistachio nuts or as a device to remove aflatoxin-contaminated product. The estimated cost of the system and the product throughput rates are comparable to the bichromatic colour sorters currently being used by industry: about $15,000 in price and a throughput of 40 nuts/s. The cost of the system could probably be reduced if it were simply added onto an existing bichromatic colour sorter. Working algorithms were developed and implemented for three specific processing streams: sinker and floater bichromatic sorter rejects and floater small shelling stock nuts. The overall minimum error rate for the system is 13% for the sinker bichromatic reject nuts, 14% for the floater bichromatic reject nuts and 15% for the small shelling stock nuts (given equal weighting to stained and unstained nuts). Since the algorithm is implemented in software, the system could be programmed to inspect other pistachio process streams or, possibly, other commodities.

Acknowledgement

Venitia Lee's assistance with the assembly code programming was an invaluable contribution to this project.

References

1 Pearson, T. C. Separating early split from normal pistachio nuts for removal of nuts contaminated on the tree with aflatoxin. Masters Thesis, University of California, Davis, CA (1993)
2 Doster, M. A. and Michailides, T. J. The development of early split pistachio nuts and their contamination by molds, aflatoxins, and insects. *California Pistachio Industry Annual Report, Crop Year 1993–1994*. Fresno, CA: California Pistachio Commission, p 68–73 (1994)
3 Doster, M. A. and Michailides, T. J. Characteristics of pistachio nuts with Aspergillus molds. *California Pistachio Industry Annual Report, Crop Year 1992–1993*. Fresno, CA: California Pistachio Commission, p. 64–68 (1993)
4 Sommer, N. F., Buchanan, J. R. and Fortlage, R. J. Relation of early splitting and tattering of pistachio nuts to aflatoxin in the orchard. *Phytopathology*, **76,** 692–694 (1986)
5 Pearson, T. C., Slaughter, D. C. and Studer, H. E. Physical properties of pistachio nuts. *Transactions of the ASAE*, **37,** 913–918 (1994)
6 Schatzki, T. F. and Pan, J. (Pers. comm.) (1994)
7 Farsaie, A., McClure, W. F. and Monroe, R. J. Design and development of and electro-optical sorter for removing BGY fluorescent pistachio nuts. *Transactions of the ASAE*, **24,** 1372–1375 (1981)
8 McClure, W. F. and Farsaie, A. Dual-wavelength fibre optic photometer measures fluorescence of aflatoxin contaminated pistachio nuts. *Transactions of the ASAE*, **23,** 204–207 (1980)
9 SAS Institute Inc. *SAS/STAT User's Guide*, Version 6, 4th Edn, Vol. 1. Cray, NC: SAS Institute Inc. (1989)