AUTOMATED SORTING OF ALMONDS WITH EMBEDDED SHELL BY LASER TRANSMITTANCE IMAGING

T. Pearson, R. Young

ABSTRACT. *Shell fragments becoming embedded in kernels during hulling and shelling is a quality control problem for the almond industry. While embedded shell is rare, with only about 0.1% of shelled kernels exhibiting this problem, the incidence has been increasing over the past several years. The industry, therefore, needs a method to remove these kernels from the process stream. A prototype device was constructed which images laser light transmitted through the kernel and automatically detects and removes kernels with embedded shell fragments. A shell fragment blocks nearly all the transmitted light, forming a dark spot in the image that is detected by a computer algorithm. The computer then activates an air valve to divert the corresponding kernel from the process stream. The sorting device has an inspection rate of approximately 40 kernels/s (100 kg/h). For a single–pass sorting operation, approximately 83% of the kernels with embedded shell were detected and removed. Additionally, 11% of the clean kernels (no embedded shell) were incorrectly classified as having embedded shell and were also removed from the process stream. Running the rejects of the first sorting pass through the system a second time recovered approximately 46% of the previously rejected clean kernels.*

Keywords. *Almonds, Image processing, Inspection, Sorters, Automation.*

The United States produces about 680 million pounds of almonds per year, accounting for over 75% of the world's supply (Anonymous, 2000). While the incidence of almonds with embedded shell has historically been quite low, at about 0.03%, it has been increasing over the past several years to as high as 0.1% in some varieties (Gray, 2001). When almonds are sold for use as food ingredients, embedded shell falls under the category of "other damage" in the U.S.D.A. standards for shelled almonds. Allowable "other damage" ranges from 2% for U.S. fancy almonds to 5% for U.S. No. 1 Pieces (Anonymous, 1997). However, confectionery manufacturers using almonds as ingredients may specify very low acceptable limits for embedded shell, sometimes as low as 1 kernel per 200 kg of bulk almonds (Gray, 2001). There is, therefore, a need for a method to remove kernels with embedded shell from the process stream. Since the color of a shell fragment is close to normal almond skin color, no commercially available machines have yet been developed that can reliably remove kernels with embedded shell. Consequently, almonds are manually inspected for embedded shell on a conveyor belt. This method is costly and at best only removes half of the affected product, since an inspector can view only one side of the passing kernel (Gray, 2001). Visual inspection must be performed in two passes to reduce levels of embedded shell to those demanded by the confectionery market.

Embedded shell occurs during the four–step shelling process (fig. 1). During the first step, the shells of the nuts are cracked as they are forced between two counter–rotating steel rollers known as a "cracking roller." Shells are removed in the second step by a "shear roller," a conveyor belt that feeds the nuts under a rotating rubber cylinder. The speed of the belt is less than the speed of the cylinder, and the shells are sheared from the kernels. In the third step, the shells and kernels are shaken on perforated tables to remove the shell pieces. Finally, the nuts are passed through an air separator to remove smaller pieces of shell and debris. The operation may be repeated six to eight times in order to remove as much of the hulls and shells as possible (Thompson et al., 1996). Small shell fragments generated by the cracking roller can become lodged in a kernel by the shear roller. In addition, shell fragments not removed by the shaking tables or air separator can be lodged in the kernel by either the cracking roller or shear roller in subsequent passes.

Incidence of embedded shell appears to be related to shell moisture content and product throughput during the shelling operation (Gray, 2001). Higher incidence of embedded shell occurs when product throughput is near or exceeding maximum for the shelling equipment and for shell moisture content above 6% d.b.

Many commercial sorters (i.e., Satake USA, Sortex, Key Technology) are available that perform sorting based on a measure of reflected light in one– or two–wavelength bands within the visible or NIR spectrum. Modern sorters collect light reflected from two or three opposing sides of the product and have product throughputs of 40 nuts per second per channel. Some sorters can have as many as 80 channels. Cameras are used as light sensors, but very little image

Article was submitted for review in May 2001; approved for publication by the Information & Electrical Technologies Division of ASAE in May 2002. Presented at the 2001 ASAE Annual Meeting as Paper No. 013099.

Reference to a company or product name does not imply approval or recommendation of that product by the U.S. Department of Agriculture or Almond Board of California to the exclusion of others that may be suitable.

The authors are **Tom C. Pearson, ASAE Member Engineer,** Agricultural Engineer, USDA–ARS, Manhattan, Kansas, and **Richard Young**, Electrical Engineer, USDA–ARS–WRRC, Albany, California. **Corresponding author:** Tom Pearson, USDA–ARS–GMPRC–ERU, 1515 College Ave., Manhattan, KS 66502; phone: 785–776–2729; fax: 785–537–5550; email: tpearson@gmprc.ksu.edu.

  ISSN 0883–8542

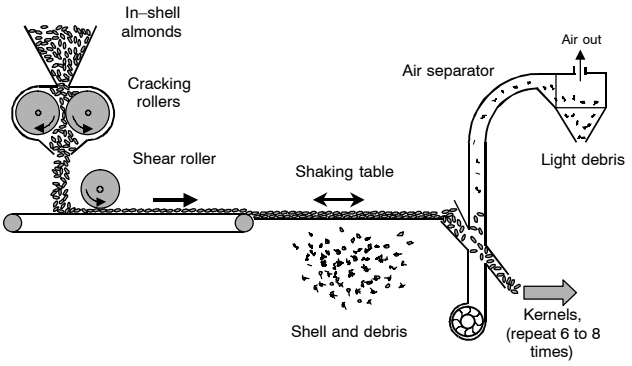

Figure 1. Schematic of a common almond shelling system. Embedded shell can occur in the cracking rollers and/or shear roller, especially when throughput exceeds system capacity and the shaking table or air separator does not remove fragments.

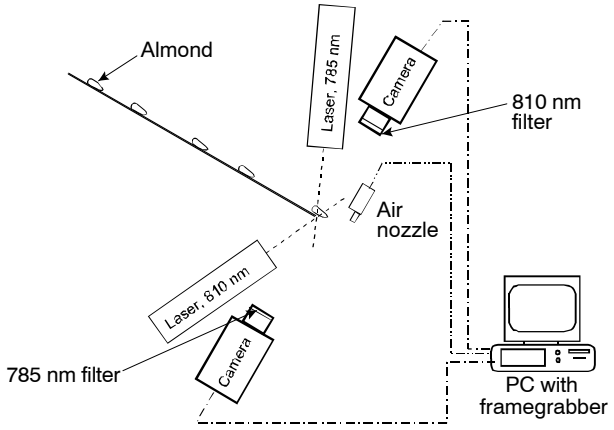

Figure 2. Schematic of laser transmittance imaging system. Note that the filters on each camera match the laser on the opposite side of the kernel so that only transmitted light is imaged.

processing is done other than to average the intensity of pixels belonging to the sorted product. These types of sorters are not effective for separating almond kernels with embedded shell (Gray, 2001).

Few developments in high–speed sorting of agricultural products have been reported in the literature. Pearson (1996) developed a machine vision–based sorting system that identified staining patterns on pistachio nuts, which indicated aflatoxin contamination. This system performed some image processing and maintained a sorting rate of 40 nuts per second per channel. Farsaie et al. (1981) developed a UV florescence–based sorting machine to detect aflatoxin–contaminated pistachio nuts. Pearson (1999b) reported the development of a sorting machine that detects internal chemical defects in almonds by measuring transmitted light through the nut. Finally, Anzai et al. (1993) patented and commercialized a transmittance–based device to sort peanuts for aflatoxin contamination. While all of these technologies have the robustness and throughput rates required to inspect almonds in a commercial setting, they are not able to detect kernels with embedded shell. One reason is that reflected light in the visible or near–infrared region of the spectrum is not useful for distinguishing shell from almond skin. Another reason is because the transmittance–based devices do not have the spatial resolution required to detect shell. However, the literature shows that imaging can be performed at the speeds needed to maintain a high throughput for sorting almonds, and transmitted light can be sensed at high speeds through a whole almond kernel.

Due to their high oil content, whole almond kernels transmit substantially more light than shell fragments in the near–infrared region of the spectrum (Pearson, 1999a; 1999b). The objective of this research was to determine the feasibility of detecting almond kernels with embedded shell by imaging light transmitted through the kernel and removing them from the main process stream.

Materials and Methods

Image Acquisition

A sorting system using two lasers and two cameras was developed to detect almond kernels with embedded shell (fig. 2). Kernels were singulated by means of a vibratory feeder (#FTO–C, FMC Corp., Homer City, Pa.) then dropped onto a chute constructed of Teflon tubing (19 mm ID, 25.4 mm OD) that was sloped at an angle of 45°. This angle was sufficient to maintain the separation between the kernels and cause them to slide while lying flat on the chute. At the end of the chute, a pair of laser beams was oriented on opposite sides of the kernel. The light transmitted through the kernel from each laser was imaged with a line–scan camera (CL–CB–0512W, Dalsa, Waterloo, Ontario).

The lasers used were 1000mW near–infrared diode lasers (HAM–850–4089, HAM–820–4143, Power Technology Inc., Little Rock, Ark.). These lasers emitted light at 785 and 810 nm, respectively. The cameras were fitted with optical band–pass filters (F10–785.0–4–1.00 and F10–810.0–4–1.00, CVI Laser Corp., Albuquerque, N. Mex.) matching the emission wavelength of the corresponding laser on the opposite side of the kernel from the camera. With this arrangement, reflected light was rejected and only light transmitted through the kernel was imaged.

The image from each camera was input to a frame–grabber board (PIXCI–D, Epix Inc., Buffalo Grove, Ill.) hosted in a PCI slot of a PC computer (Pentium III, 733 MHz with ASUS CUBX motherboard). The frame grabber supplied camera clock signals. The kernels were sliding at a speed of about 2 m/s as they left the chute, and it was found, by trial and error, that an 8–kHz line–scan rate maintained the appropriate aspect ratio of the kernel in the resulting image. The frame–grabber image–capturing settings were loaded onto the frame grabber using the C programming language and functions from the 32–bit DOS XCOBJ and PXIPL libraries (Epix Inc., Buffalo Grove, Ill.). Frames, comprising five image lines, were continuously stored on the frame–grabber buffer memory, then transferred to the PC memory. When no kernel was present, the image was almost completely dark and all pixels had intensities less than 10. The presence of a kernel was detected when 20 or more pixels within one frame had an intensity greater than 40. When this occurred, the current frame was stored, as well as the next 15 frames, to form an image of the kernel. The cameras were set to average two adjacent pixels so that the output of each camera was 256 pixels wide. The images output by each camera were combined on the frame–grabber board to form a single image containing the captured image from both sides of the kernel (512 pixels wide by 80 pixels tall). Images of 300 kernels with embedded shell and 300 normal kernels were captured in this manner for off–line analysis. Mission

variety almonds were used as this variety is one of the most favored for use in confectionery products, and they are known to have moderate levels of embedded shell.

Image Feature Extraction

A 3 × 3 minimum filter was used to eliminate the effect of light diffracting around the edges of the almond kernel and causing camera saturation. In addition, this filter tended to enhance the contrast between shell fragments and the kernel in the digitized image (fig. 3). It was desirable to find a small number of features that could be extracted from the digitized images in real time that would distinguish kernels with embedded shell from normal kernels. Two different types of features were extracted from the images. The first was the number of pixels that fell within a "valley" in the image intensity map. The second feature type was two–dimensional histogram bin values based on the image intensity and gradient.

As can be seen in figure 4, a deep "valley" in a one–dimensional profile plot along a row or column in an image gives a strong indication of the presence of an embedded shell fragment. For each image, profile plots for every row and column were analyzed. The number of pixels from these profile plots representing valley regions was tallied, giving a grand total for each image. It is, therefore, possible that pixels could be counted twice, if criteria for a valley were met in both horizontal and vertical directions. A region was considered a valley using a variety of criteria based on gradient, pixel intensities, and width of the potential valley. For each set of parameters, the number of pixels were counted and stored in a data set. The following were the parameters for considering a region a valley:

1. Edges of the valley needed to have a downward and upward gradient greater than a specified threshold. The five different threshold values used were 8, 12, 16, 20, and 24 pixel intensity levels. Upward and downward gradients were computed using equations 1 through 4.

$$\text{upward gradient (horizontal direction)} = I_{x+1,y} - I_{x-1,y} \quad (1)$$

$$\text{upward gradient (vertical direction)} = I_{x,y+1} - I_{x,y-1} \quad (2)$$

$$\text{downward gradient (horizontal direction)} = I_{x-1,y} - I_{x+1,y} \quad (3)$$

$$\text{downward gradient (vertical direction)} = I_{x,y-1} - I_{x,y+1} \quad (4)$$

where $I_{x,y}$ is the pixel intensity at location x, y, in pixels, from the upper left corner of the image.

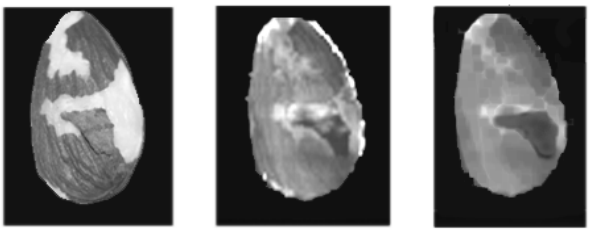

Figure 3. Grayscale reflectance image of an almond with embedded shell (note that the contrast between the kernel and embedded shell is very low), raw laser transmittance image (note light diffracting around edges of kernel), 3 × 3 minimum filtered image showing a high level of contrast between kernel, and embedded shell and effect of diffracting light around edges is removed (left to right).

2. All pixels within a valley region had to have an intensity less than a specified threshold. The seven threshold values were 80, 100, 120, 140, 160, 180, and 200.

3. Valley size was restricted to less than a specified threshold. Threshold values were 15, 20, 25, 30, 35, and 40 pixels wide.

Pixel counts for each image were computed with all combinations of the above parameters. A total of (5 × 7 × 6) 210 values resulted, each of which was used as an image feature.

As can be seen from figure 4, a sharp slope characterizes edges of a shell fragment with a moderate to high pixel intensity. It is in contrast to the edges of the kernel against the background, which also have high slopes but low pixel intensities. In addition, the shell regions of the image are characterized by moderately low intensities and low gradients. A two–dimensional histogram was generated based on both gradient image and intensity image, and comprised gradients in both horizontal and vertical directions. Gradients for each direction were computed using equations 5 and 6.

$$\text{gradient (horizontal direction)} = |I_{x+1,y} - I_{x-1,y}| \quad (5)$$

$$\text{gradient (vertical direction)} = |I_{x,y+1} - I_{x,y-1}| \quad (6)$$

Each bin of the histogram comprised the sum of pixels within an image that had both a gradient and intensity greater than specified threshold values. Gradient threshold values ranged between 0 to 24 in steps of 2 for a total of 13 levels. Pixel–intensity threshold values ranged from 0 to 130 in steps of 5 for a total of 27 levels. Thus, there were a total of (13×27) 351 histogram bins as potential distinguishing features. These 351 features were in addition to the 210 features computed earlier, giving a grand total of 571 features to be evaluated.

Feature Selection

In high–speed sorting operations, only a few discriminating features can be economically computed in real time to

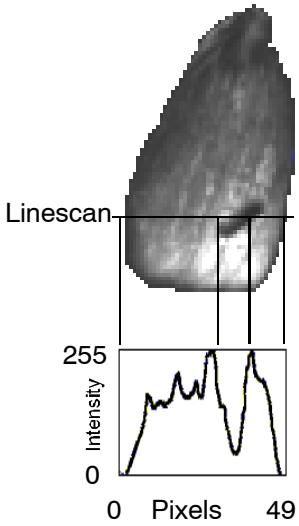

Figure 4. Profile plot across image with embedded shell. Note that a "deep valley" with high pixel intensities on the edges of the shell characterizes the portion of the profile plot corresponding to shell pixels.

maintain a sorting rate of 40 nuts per second. The feature selection goal was to find the best single feature, combination of two features, or combination of three features out of the 571 computed features that was best able to distinguish kernels with embedded shell from normal kernels. To select a small number of features from a large data set where high levels of multi–colinearity exist, stepwise methods are not very optimal (Huberty, 1994). The procedure used for this study performed an exhaustive search using all possible combinations of one, two, or three features. The data set was divided into two groups, a training set and a validation set. The training set comprised the odd–numbered kernels in the sequence they were sampled while the a validation set contained the even–numbered samples. Discriminant analysis was used as the classification procedure, using both pooled and non–pooled covariance matrices (Huberty, 1994). The Mahalanobis distances were computed from each kernel to the embedded shell positive and negative groups. A kernel was classified into the group with the lowest corresponding Mahalanobis distance. Sample means and co–variance matrices for each group were computed using the training set. The combination of one, two, or three features that obtained the lowest classification error rate on the validation set was selected. These features were then used for distinguishing kernels with and without embedded shell during real–time sorting.

System Testing

After features for separating kernels with and without embedded shell were selected, the PC was programmed to extract these features in real time and activate the air nozzle if the kernel was classified as having embedded shell. The 3× 3 minimum filtering was performed immediately after a frame was transferred to the PC and was accomplished before the next frame was ready to be transferred. In addition, gradients in the horizontal (x) direction were computed and pixel counts associated with the horizontal gradient were made between frame transfers. It greatly reduced the computation time required after the complete image was captured. Vertical gradients and associated counts were made after the complete image was captured. This process took approximately 0.6 ms. If a kernel was classified as having embedded shell, a digital output board (KPCI–3140, Keithley Instruments, Inc., Cleveland, Ohio) was programmed to output a pulse for 4 ms, which activated an air valve (35A–AAA–DDBA–1BA, Mac Valves, Inc., Wixom, Mich.) to divert the kernel. The valve was activated immediately upon determination that a kernel was to be rejected. By activating the air valve immediately, the kernel would have only traveled approximately 1 to 2mm during the 0.6–ms lag time after the kernel left the field of view of the camera. This short lag time alleviated any need to monitor the location of the nut and control the air valve activation time. One thousand almonds with embedded shell, as well as 1000 normal almonds, were tested and the classification accuracy was recorded. Rejects from the first pass were run through the sorting system a second time to recover normal kernels incorrectly classified as having embedded shell. Kernels were processed at a rate of approximately 40 per s.

Results and Discussion

Feature Selection

It was found that a three–feature, linear (pooled covariance matrices) discriminant function gave the best classification accuracy on the validation set. Each of the three features chosen were two–dimensional histogram bin values with the following criteria:

1. number of pixels with both a gradient greater than 4 and intensity greater than 120
2. number of pixels with both a gradient greater than 8 and intensity greater than 110
3. number of pixels with both a gradient greater than 18 and intensity greater than 55

Using these criteria, the validation set had a false–positive classification error rate of 11% and a false–negative classification error rate of 14%. These features indicated that the pixels surrounding the edge of a shell fragment were most useful for identifying kernels with embedded shell. Valley features, rather than pixel counts, were selected as the optimal features for the one and two feature discriminant functions. However, classification results using one or two features resulted in error rates roughly twice as high as the three–feature discriminant function selected.

System Testing

The overall sorting error rate was 18.2% for the kernels having embedded shell and 11.5% for the normal kernels, after running the kernels through the sorting system one time. The increased error rate compared with the error rate found when performing the feature selection was primarily due to kernels touching each other as they were fed and/or were simply not diverted to the correct stream, even though the classification was correct. Running the rejects of the first pass through the sorting system a second time resulted in 45.6% of the normal kernels and 8.0% of the kernels with embedded shell being accepted. Using a two–pass sorting operation, then, where rejects of a first pass were processed a second time and the accepts of the first and second passes were blended together, resulted in 75% of the original embedded shell kernels and 6.2% of the original normal kernels being rejected. A two–pass sorting operation applied to kernels containing 0.1% embedded shell would result in an accept stream containing only 0.025% embedded shell and a reject stream comprising 6.2% of the original product.

Conclusion

A sorting system was developed to separate almonds having embedded shell from normal kernels. The system utilized light emitted by near–infrared lasers and transmitted through the kernel. Line–scan cameras to inspect both sides of the kernels simultaneously imaged the transmitted light. The inspection algorithm consisted of extracting three image features, which can be computed in real time at rates of up to 40 kernels per s. Each of the three features consisted of pixel counts where pixel intensity and gradient exceeded predetermined levels. Classification was performed using a linear discriminant function. Testing of the sorting system showed that 88.5% of the normal almonds and 82% of the almonds with embedded shell were correctly classified after one sorting pass. Running the rejects through the sorter a second time and blending second–pass accepts with first–pass

accepts grouped 75% of the original embedded–shell kernels with 6.2% of the original normal kernels.

Acknowledgements

The Almond Board of California provided partial funding for this project.

References

Anonymous. 2000. Almond Almanac 2000. Almond Board of California. Modesto, Calif.

Anonymous. 1997. United States standards for grades of shelled almonds. United States Department of Agriculture – Agricultural Marketing Service – Fruit and Vegetable Division – Fresh Products Branch, Washington D.C.

Anzai, K., A. Shibayama, S. Hirano, and H. Sasaki. 1993. Sorting apparatus utilizing transmitted light. U.S. Patent 5,190,163.

Farsaie, A., W. F. McClure, and R. J. Monroe. 1981. Design and development of and electro–optical sorter for removing BGY fluorescent pistachio nuts. *Transactions of the ASAE* 24(5): 1372–1375.

Gray, G. 2001. Personal communication. Blue Diamond Growers, Sacramento, Calif.

Huberty, C. J. 1994. *Applied Discriminant Analysis*. New York: John Wiley and Sons.

Pearson, T. C. 1996. Machine vision system for sorting stained pistachio nuts. *Lebensmittel–Wssenschaft und –Technologie* 29(3): 203–209.

_____. 1999a. Spectral properties and effect of drying temperature on almonds with concealed damage. *Lebensmittel–Wissenschaft und –Technologie* 32(2): 67–72.

_____. 1999b. Use of near–infrared light transmission to automatically detect almonds with concealed damage. *Lebensmittel–Wissenschaft und –Technologie* 32(2): 73–78.

_____. 2001. Detection of pistachio nuts with closed shells using impact acoustics. *Applied Engineering in Agriculture* 17(2): 249–253.

Thompson, J. F., T. R. Rumsey, and J. H. Connell. 1996. Drying, hulling, and shelling. In *Almond Production Manual*, ed. W. C. Micke. Oakland, Calif.: University of California, Division of Agriculture and Natural Resources.