J. King Saud Univ., Vol. 12, *Agric. Sci.* (1), pp. 69-79, Riyadh (1420/2000)

AGRICULTURAL ENGINEERING

Date Inspection by Color Machine Vision

Abdulrahman A. Al - Janobi
Department of Agricultural Engineering, College of Agriculture
King Saud University, Riyadh, Saudi Arabia

(Received 12/2/1419; accepted for publication 30/6/1419)

Abstract. A color computer vision system consisting of a microcomputer with an image frame grabber and a CCD color camera was developed for sorting and grading Saudian dates (Sifri variety) based on a color thresholding technique. Samples of dates graded by experts were used for the experiment. Histograms of the RGB color bands were tested. Histogram features mean, variance, and wrinkle ratio were used as the criteria for classification of dates into four quality classes. The system showed an average classification error of 1.8% using features from the red color band.

Introduction

Dates are one of the oldest fruits and considered as the traditional agricultural product in the Arab world. In Saudi Arabia, dates are highly regarded and widely consumed both as a fresh fruit and in different food forms. Dates are thought of as highly nutritious food and contain all the basic elements required for a balanced diet. Saudis are the world's highest consumers, with an average annual per capita consumption of 40 kg [1, pp. 33-38].

Saudi Arabia is one of the largest producers of dates in the world, with total crop yield estimated to be about 560,000 tonne of dates per year [2]. Here, the date industry is considered to be of increasing importance because of the growing demand for the supply of good quality dates and date pastries. The huge quantity of dates produced is processed in a number of dates processing plants. There are more than 22 large date processing plants operating in Saudi Arabia for grading and packaging dates [1]. The criteria for grading dates in these plants are based on visual judgment. Processing steps before washing such as harvesting, field grading, transportation to plants, and grading are called preprocessing. Processing includes the steps from washing to packaging.

Date processing plants have been established in Saudi Arabia to facilitate processing steps and to produce a high-quality product. The processing steps are as follows:

1. Removal of foreign objects such as branches and small stones using rotating sieves,
2. Manual separation of dates with surface defects and/or overdried fruit,
3. Hot water spraying (60° C), followed by washing with warm water to clean dates,
4. Drying of dates with hot air (70° C),
5. Cooling of dates to 20° C,
6. Manual classification of dates into quality classes.

These processing steps are mechanized, except for steps 2 and 6. Inspection of dates before and after washing is carried out manually. Manual grading of dates has certain disadvantages. Results may fluctuate during the workday as worker concentration varies. Cost is high and the labor supply is unstable, particularly in Saudi Arabia where most of the laborers employed are from outside of the country. Because of the disadvantage of manual grading, the need to automate the grading process is increasing significantly.

The existing packing and processing plants in Saudi Arabia utilize less than 6% of the total annual date production [3]. However, economic success of date processing plants will eventually open international markets for this highly valued natural food. Standardization of quality will be an essential need. Consequently, efficient inspection and grading systems will be of significant value.

Previous Work

Researchers have reported a number of methods to sort dates into quality classes. Huxsoll and Reznik [4] used mechanical methods. Their system required that individual dates slide down a 1.5m tube inclined at 40 to 50° from horizontal, onto a metal impingement plate. Differences in sliding velocity were related to variations in surface friction. Impingement reaction allowed dry dates to travel farther than soft dates and fall into different channels.

Chesson *et al.* [5] developed a vacuum system for separating dates. The system contained a press wheel and a drum with three zones of variable surface vacuum. The system was capable of separating freshly harvested dates into three moisture grades, and separating 98% of high-moisture fruit from fresh dates. The system was unable to meet industry grade requirements.

The difficulty in grading dates is related to the range of grading criteria. There are hundreds of date fruit varieties which vary in texture, color, moisture content, sugar content, size, and shape. Date fruits can basically be categorized into three types according to moisture content. They are soft, semi-dry, and dry. In the date industry, grading is based on color, size, surface defects, and texture. Color is an important factor in distinguishing between acceptable date fruits and damaged or immature dates. The color of acceptable dates is relatively uniform and predominantly light amber in color. Size is affected by variety and the condition of the producing trees. Dates are rejected if they are significantly larger or smaller than the subjective average size of the dates. Texture is a useful factor to identify overdried "hard" dates. Uniformity of shape is also an important factor in identifying overdried dates and dates with defects.

The current grading procedure in date industries is to present dates on a belt conveyor. Grades are determined by a number of industry trained inspectors positioning along the belt conveyor. This manual grading is very subjective, monotonous, inconsistent, and labor intensive. Dates are also highly variable and difficult to evaluate. Therefore, an automated date grading system is required to evaluate date quality objectively and assign the correct grade.

Early in the development of digital computers, researchers have attempted to design machines with "vision" capability. Coupling a video camera with a specialized computer has enabled the development of machines capable of visual interpretation. Such machine vision systems have been accepted in food and agricultural industries, and many systems have been installed for inspection and grading of various food and agricultural products [6].

The earliest work on sorting dates with machine vision was done in 1986 by VARTEC Co. in California. The algorithms in this project were developed to estimate moisture content from surface texture. Unfortunately, this work is undocumented and no longer active [7].

Wulfsohn *et al.* [8] studied the use of image processing techniques on two date varieties (Majhul & Zahidi). Their work addressed separating "good" dates from "defective" dates using thresholding techniques. A color camera was used to capture date images. They found that specular reflectance (glare) was a major problem. Relative reflectance was measured in the range of 400 to 1,100 nm for good and defective dates. The largest difference between dry dates and dates with blister defects occurred at 600 nm for Majhuls and in the 450 - 600 nm range for the Zahidi variety. An infrared cutoff filter was then used for obtaining images of both varieties, and an infrared cutoff filter combined with an optical filter for Majhul. They noted that the red band image was most effective for detecting defective Majhul dates. The green band image performed best for Zahidi dates.

Al-Janobi [7] developed image processing techniques to grade date fruits (Deglet Noor variety) into quality classes based on color and texture analysis. The highest classification accuracy was 98.4%. Classification accuracy of an accompanying monochrome system was 77%. Extended processing time limited the classification rate to only 14 fruits per minute.

The overall objective of the current project was to investigate the use of color computer vision for grading Saudian dates (Sifri variety). Specific objectives were to:

1. Develop image processing techniques to grade dates into quality classes,
2. Evaluate the accuracy of the proposed algorithm by comparison with manual inspection.

Materials and Methods

Computer vision system

A block diagram of the computer vision system is shown in Fig. 1. The system consisted of an Everex microcomputer (486/33) equipped with a Data Translation DT2871 color image frame grabber, a Sony XC711 CCD color camera, a Sony PVM1342Q monitor, and a lighting system consisting of four 40W incandescent lamps and a diffuser of approximately 20cm diameter and 15cm height prepared from a PVC pipe with the inner surface painted white. The color camera captured the image and sent red, green, and blue (RGB) video output signals to the color image frame grabber. The frame grabber digitized the RGB analog video signals to three digital images (512 x 512 pixels) and stored them in three on-board buffers. The gray level of each pixel ranged from 0 to 255. The acquired digital images were then transferred to the disk storage of the microcomputer and to be retrieved later for analysis. A menu-driven image processing software package, Global Lab Color [9], was used for both acquiring and analyzing images of date samples.

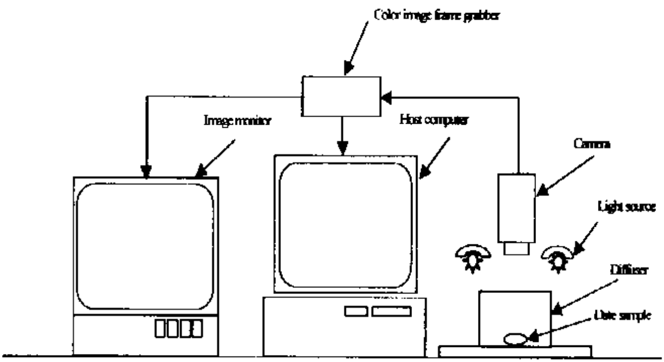

Fig. 1 Machine vision setup for inspection of dates.

Procedures

Dates in Saudian plants are separated into two classes: acceptable and rejected. Acceptable dates with wrinkled skin are typically damaged during washing steps. Therefore, separating acceptable dates into two classes is needed. Acceptable dates can be classified as A and B, where class A is dates with no surface wrinkles and class B is dates with few wrinkles. Rejected dates contain the defective and the overdried dates. The rejected dates can be classified as class C for defective dates and class D for overdried dates. Speckles caused by insects, cuts by bird pecking, and other defects like bruises and some rots are present on the defective dates, whereas the overdried dates have heavy wrinkles and they are present on about two-thirds of the date surface.

From dates graded by experts, a total of 400 samples, 100 from each of the four classes, were randomly selected. The histogram of the RGB bands representing the relative frequency of occurrence of the various gray levels in the images of these samples was tested. Figure 2 shows the average red-band histogram of all the samples in each quality class. The intensity values show almost the same pattern range for all the classes except class D. Unlike classes A, B, and C, the pattern for class D shows a trend of increasing with respect to intensity in the high range. However, the transformation ranges distinguishing different regions in the date image were obtained from studying the date images. From this, the pixel value of the background was found in the range 0 to 36, whereas for the surface and wrinkle of date, the pixel values were found approximately in the ranges 37 to 220 and 221 to 255, respectively. These transformations permitted segmentation of certain gray level regions representing different features of the image. The histogram can be separated into three parts, with the following ranges of gray level:

1. Background color, 0 to 36,
2. Surface color, 37 to 220,
3. Wrinkle color, 221 to 255.

Amplitude thresholding [10] was used to segment the date sample from the background. First, the histogram of the image was examined for peaks, valleys and planes. The low intensities were then thresholded to isolate the background. Finally, the remaining intensities were thresholded to isolate the date sample from the background.

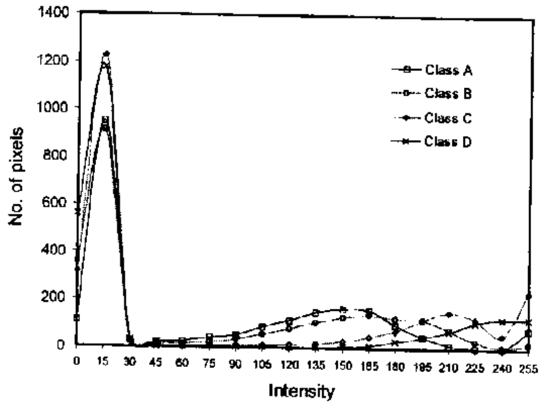

Fig. 2. Average red-band histogram of date samples in four classes.

Histogram features consisting of the mean, variance, and wrinkle ratio were measured as the criteria for date classification. These features were measured after segmenting the date object from the background. The following equations describe the measurement of these criteria:

$$\text{Mean (m)} = \sum_{x=37}^{x=255} x\, p(x) \tag{1}$$

$$\text{Variance (v)} = \sum_{x=37}^{x=255} (x-m)^2\, p(x) \tag{2}$$

$$\text{Wrinkle ratio (w)} = \frac{\sum_{x=221}^{x=255} x\, p(x)}{\sum_{x=37}^{x=220} x\, p(x)} \tag{3}$$

Where:

x = gray level ranging from 0 to 255

$p(x)$ = ratio of the number of pixels with gray level 'x' to the total number of pixels in the region

These three features were measured for each color RGB band, totaling twelve measurements for all the date samples in the four quality classes. Table 1 summarizes the measurements of the features extracted from the date images in the red color band.

Table 1. Measurement of gray level features of date samples in red color band

Criteria	Minimum	Maximum	Average	STD*
		Class A		
Mean	51.92	92.72	71.92	13.01
Variance	88.83	2024.34	1069.39	568.90
Wrinkle ratio	0.0	0.03	0.01	0.01
		Class B		
Mean	65.17	93.71	78.86	8.52
Variance	0.07	0.20	0.13	0.04
Wrinkle ratio	0.0	0.08	0.04	0.02
		Class C		
Mean	43.45	118.51	81.10	21.76
Variance	0.05	0.52	0.29	0.15
Wrinkle ratio	0.01	0.44	0.23	0.13
		Class D		
Mean	74.30	118.52	95.32	13.66
Variance	0.051	13.56	7.19	3.74
Wrinkle ratio	0.48	0.79	0.63	0.09

* Standard deviation

Results and Discussion

Discriminant analysis was conducted using the SAS software package [11]. This program was used to classify date samples into four quality classes using the gray level features. The analyses were carried out using normal density estimates with unequal variance. This method was used because of normally distributed feature data. The features extracted from the three color bands were used independently and in certain combination in the classification procedure.

Tables 2-4 show the results of the percentage classification error of 100 date samples in each of the four quality classes for the green, blue, and red color bands with the three gray-level features. They show that classification of date samples with the features from the green and blue color bands (Tables 2 and 3) was not accurate compared to the classification with the same features from red color band (Table 4). The minimum average errors in classifying date samples based on the green and blue color bands were 17.8% and 26.8% respectively, whereas it was only 1.8% in red color band. Therefore, the procedure using the gray level features from the red color band was accepted for classifying date samples.

Table 2. Percentage classification error of date samples into four quality classes with three gray-level features (green color band)

Features	Quality classes and their classification errors				Average
	A	B	C	D	error
m	83	37	37	19	55.8
v	11	30	30	0	23.3
w	97	92	92	35	72.5
m-v	12	19	19	1	17.8
m-w	86	37	37	24	56.5
v-w	11	29	29	0	23.3
m-v-w	12	19	19	1	17.8

m- mean
v- variance
w- wrinkle ratio

Table 3. Percentage classification error of date samples into four quality classes with three gray-level features (blue color band)

Features	Quality classes and their classification errors				Average
	A	B	C	D	error
m	60	68	70	28	56.5
v	38	51	34	0	30.8
w	62	57	89	71	69.8
m-v	37	42	28	0	26.8
m-w	59	68	71	29	56.8
v-w	36	52	33	0	30.3
m-v-w	34	50	24	0	27.0

m- mean
v- variance
w- wrinkle ratio

Table 4 shows that with the 'mean' feature, the date samples in all the four classes were highly misclassified with an average error of 54.3%. The second highest misclassification occurred with the 'wrinkle ratio' feature, giving an average classification error of 17.0%. However, the samples in D were classified correctly. With the 'variance' feature, samples in classes A and B were classified correctly, whereas 3% of the samples in class D and 34% of the samples in class C were misclassified, with an average error of 9.3%. This result shows that the classification, on an average basis, was more accurate with the 'variance' feature than with the other two features when they were used independently.

Table 4. Percentage classification error of date samples into four quality classes with their gray-level features (red color band)

Features	Quality classes and their classification errors				Average
	A	B	C	D	error
m	53	42	88	34	54.3
v	0	0	34	3	9.3
w	17	35	16	0	17.0
m-v	0	7	16	3	6.5
m-w	13	38	9	0	15.0
v-w	0	79	0	0	19.8
m-v-w	0	6	1	0	1.8

m- mean
v- variance
w- wrinkle ratio

Table 4 also shows the results of classification with the three features used in combination. Based on average percentage classification error, the feature combination 'variance and wrinkle ratio' performed poorly (average classification error of 19.8%). However, the samples in classes A, C, and D were classified correctly. Comparing the different feature combinations, the 'mean, variance, and wrinkle ratio' combination gave more accurate classification, showing an average classification error of only 1.8%. With this feature combination, only 6% of the samples in class B and 1% of the samples in class C were misclassified. All samples in both A and D classes were classified correctly.

Table 5 details results of classification and average error for date samples in the four quality classes with the three features combined (mean, variance, and wrinkle ratio) for the red color band. It shows the exact locations of the misclassified date samples. From classes B and C, the number of samples misclassified was only 6% and 1% respectively, whereas all the samples from classes A and D were classified correctly. The misclassified samples from the B and C classes were interchanged (Table 5). This effect shows that the procedure resulted in only the samples in classes A and D being classified correctly with feature combination 'mean, variance, and wrinkle ratio'. However, this model shows its capability to classify dates (Sifri variety) with minimum classification error for each class. Classification error for each class was less than 10%, which is the acceptable manual classification error limit in the date industry.

Table 5. Classification of date samples and the percentage error with three gray-level features combined (red color band)

Classes (to) → (from) ↓	A	B	C	D	Total
A (n-100)	100	0	0	0	100
B (n-100)	0	94	6	0	100
C (n-100)	0	1	99	0	100
D (n-100)	0	0	0	100	100
Total	100	95	105	100	400
Error, %	0	6	1	0	1.8

n- number of samples

Summary and Conclusion

Histogram features based on the amplitude thresholding technique were used for grading Saudian dates (Sifri variety) into four quality classes. The features extracted from the red, green, and blue color bands were tested in the classification procedure. The features used independently and in different combinations for the red color band showed comparatively good results. The 'mean, variance, and wrinkle ratio' combination provided highly accurate results, with an average classification error of 1.8%.

References

[1] Sawaya, N. W. *Dates of Saudi Arabia*. Riyadh: Safir Press, 1986

[2] Food and Agriculture Organization of the United Nations. *FAO Production Yearbook, 1993*. Rome, 1994.

[3] Mekky, M.S. *Saudi Dates Handling and Processing in Improving its Marketing Efficiency*. Al – Hassa, Saudi Arabia: Ministry of Agriculture and Water, Regional Agricultural Research Center, 1989.

[4] Huxsoll, C.C. and Reznik, D. "Sorting and Processing Mechanically Harvested Dates." *Date Grower's Institute Report*, 46 (1969), 8-10.

[5] Chesson, J. H., Burkner, P.F. and Perkins, R.M. "An Experimental Vacuum Separator for Dates." *Transactions of the ASAE*, 22 (1979),16-20.

[6] Kranzler, G. "Applying Digital Image Processing in Agriculture." *Agr. Eng.* 66, No. 3 (1985), 11-15.

[7] Al-Janobi, A. *Machine Vision Inspection of Date Fruits*. Ph.D. Thesis. Oklahoma State University, Stillwater, 1993.

[8] Wulfsohn, D., Sarig, Y. and Algazi, R.V. "Preliminary Investigation to Identify Parameters for Sorting of Dates by Image Processing." *ASAE Paper No. 89-6610*, St. Joseph, MI:ASAE, 1989.

[9] Data Translation, Global Lab Color. Data Translation Inc., 100 Locke Drive, Marlboro, MA, 1990.

[10] Jain, A. K. *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ: Prentice Hall, 1989

[11] SAS/STAT *User's Guide*. Release 6.03. Cary, NC: SAS Institute Inc., 1988.

فرز التمور بواسطة آلة الرؤية الملونة

عبدالرحمن عبدالعزيز الجنوبي

جامعة الملك سعود، كلية الزراعة، قسم الهندسة الزراعية

(قدم للنشر في ١٤١٩/٢/١٢هـ؛ وقبل للنشر في ١٤١٩/٦/٣٠هـ)

ملخص البحث. تعتبر المملكة العربية السعودية من أكبر دول العالم في إنتاج التمور، حيث تنتج أكثر من ٥٦٠ ألف طن سنويا. تُعالج هذه الكمية الضخمة من التمور في أكثر من ٢٢ مصنعا. تعتمد معايير فرز التمور في هذه المصانع على الرؤية الظاهرية. ولكون هذه العملية تتم يدويا، لذلك فهي مكلفة ماديا، وتستغرق وقتا، ولا يعطى الفرز اليدوي توازنا في الفرز خلال فترة العمل. لذلك فإن الحاجة ملحة إلى إدخال ميكنة عملية الفرز. في هذا العمل، اختيرت آلة الرؤية الملونة لفرز وتصنيف صنف من التمور السعودية (صنف الصفري) اعتمادا على تقنية التصنيف الملونة (Thresholding). يتكون النظام من حاسب آلي مع كرت التقاط الصور، وآلة تصوير ملونة. وكانت نسبة الخطأ في الفرز حوالي ١,٨٪ بناء على الطيف الأحمر.