DATE VARIETY RECOGNITION AND SUGAR CONTENT ESTIMATION VIA COLOR ANALYSIS

Moustafa A. Fadel [*], Kurmestegy, L., Rashed, M., and Rashed, Z

***United Arab Emirates University, Faculty of Agricultural Sciences, Plant Production Department**

ABSTRACT

Lolo, Khalas, Berhi, Fard and Bomaan are among the most famous date varieties in the UAE. Color properties of those varieties were investigated to explore the possibilities of using them as a segregation criterion. Date image was taken using a digital camera and the color was fragmented into Red, Green and Blue (R, G, B). Using image analysis software, frequency distribution of image pixels colors was calculated and the mean value was used to compare between the different varieties. Results showed that, color luminosity and all color ingredients (R, G and B) may be used to differentiate between Lolo and Bomaan where the "blue" should be used in order to separate between Berhi and Bomaan as well as between Berhi and Khalas. On the other hand, Fructose, Glucose and Sucrose content were determined for each variety using the chromatograph. A prediction equation of sugar content as a function of color intensity was developed for each date variety.

This research put the foundation for a new computerized date variety recognition system. Furthermore, results pave the way to initiate a novel date sugar content estimation technique

INTRODUCTION

Fruit's color is an important property, which may be used to demonstrate its readiness for harvesting and to qualify it marketwise. Machine vision can be used to enable quantity measurements for qualitative criteria. It is based on different techniques such as image processing and pattern recognition etc., .

Few research papers focused on applying machine vision technique on dates. Most of them studied each variety solely. It was necessary to study different varieties of dates to lay the foundation for a machine

vision system, which has the capability to differentiate between various date varieties as well as estimating sugar content of each variety

Review of Literature

Al-Janobi (1998) Applied the line-scan based vision for inspecting fast moving date fruits, where it is capable of determining the color/quality of date fruits

Adamsen *et al*, (1999) used the green to red (G/R) ratio for each pixel in an image for cropped images from a digital camera to determine the effects of elevated CO_2 and limited nitrogen on spring wheat.

Williams *et al* (1996) used a machine vision system to determine ripeness and harvest ability of peanut crop.

Al-Hooti and Sidhu (1997) stated that, because of significant international trade in date fruits, the need for objective color measurement for maintaining strict quality standards has become obvious.

They used a Macbeth color checker spectrophotometer to quantify color of date fruits to compare between different date cultivars and for quality control of processed date products in the international trade.

Davies and Perkins (1991) concluded that a combination of cool white and daylight fluorescent tubes produces the sharpest contrast between different grades of fruit.

Al-Janobi (2000) developed a color computer vision system consisting of a microcomputer with an image frame grabber and a CCD color camera for sorting and grading Saudian dates based on color threshold technique.

Al-Janobi (2000) stated that 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.

Wulfshon et al., (1989) used the a color camera to capture date fruit images to determine the relative reflectance in the range of 400-1000 nm for good and defective dates. Furthermore, they used an infrared cutoff filter. They noted that the red band image was most effective for detecting defective Majhul dates, the green band image performed best for Zahidi dates.

MATERIALS AND METHODS

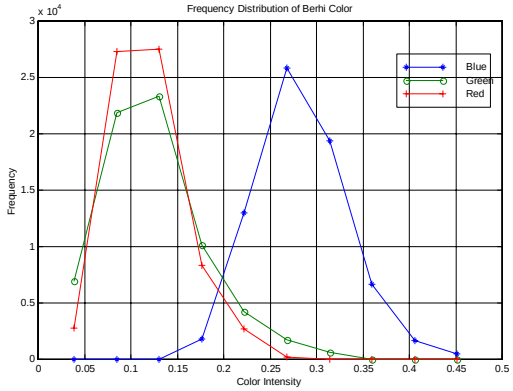
The five varieties Lolo, Bomaan, Khalas, and Fard are commonly grown in the Gulf area. Ten samples of each variety were studied. An image of each sample was captured using Sony Mavica digital camera. Image resolution was 1912 x 916 pixels. Fluorescent was used as illumination source as recommended by referenced papers. A brief code was developed in order to analyze color of each image separately and plot the RGB frequency distribution (Color Histogram) . For each variety, image data for the whole ten samples were stacked together to determine its RGB frequency distribution. Meanwhile, standard deviation, mean and median were calculated for each sample as well as for the stacked data.

Analysis of variance (ANOVA) between means was used as a statistical tool to find out if there was a significant difference within the same variety as well as between varieties. On the other hand, Sucrose, Fructose, and Glucose content in each sample were determined using the Chromatograph.

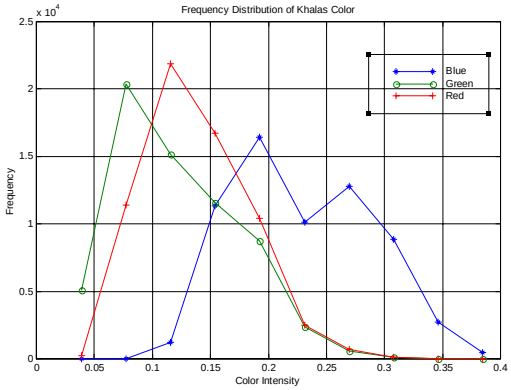
Each sugar content was correlated to the color intensity mean of R, G, and B resulted from the color analysis process of the sample's image. A prediction equation was concluded to estimate the content of each sugar type in each variety.

RESULTS AND DISCUSSIONS

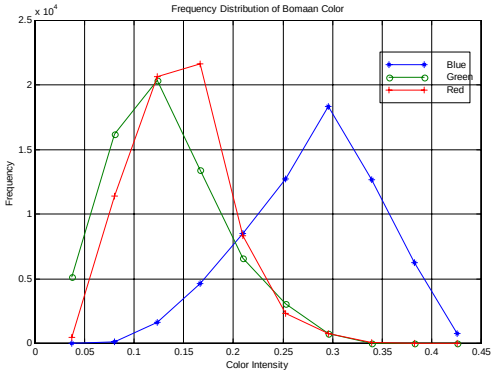
Figure 1 shows the frequency distribution of each variety. Each color component (Red, green and Blue) is represented as a relation between color intensity and the number of pixels having the same value. It is easy to conclude that for Berhi, Bomaan and Khalas varieties, the blue color median is considerably deviated from those of the green and red colors. On the other hand, for Fard and Lolo, all color components (Red, green and Blue) are squeezed in a narrow strip of the x-axis.



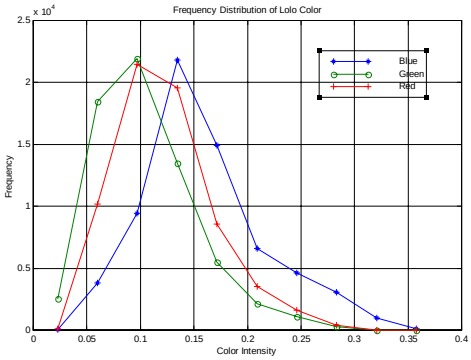
(a)



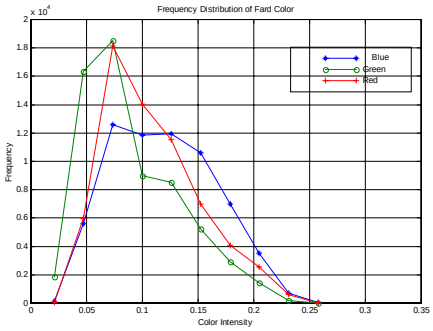
(d)



(b)



(e)

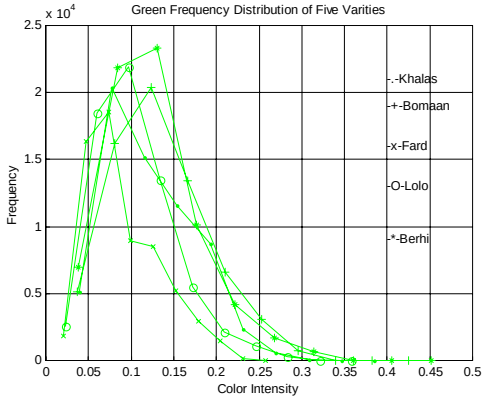


(c)

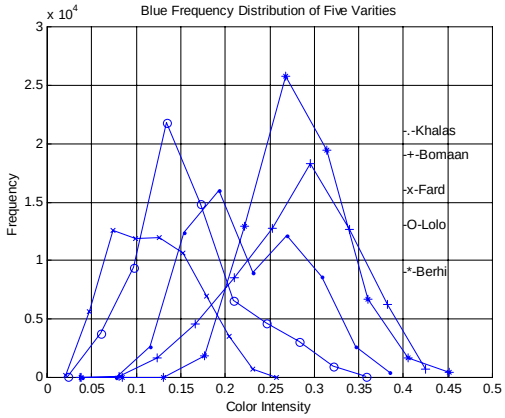
Figure 1: Color Frequency distribution of the five dates varieties (a, b, c, d, e).

Table 1 Color components analysis for date varieties

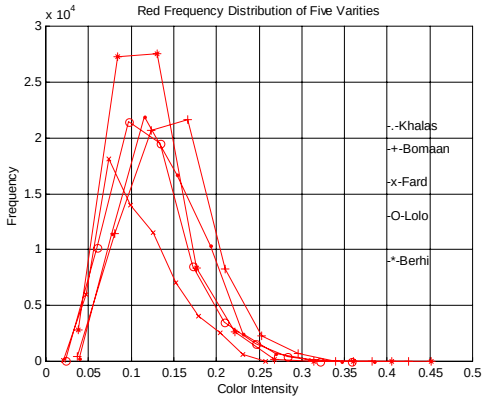
	Red	Green	Blue
Fard			
Mean	0.1087	0.0908	0.1194
Median	0.102	0.0784	0.1137
Stdv	0.0417	0.0427	0.0448
Khalas			
Mean	0.1378	0.121	0.2302
Median	0.1333	0.1137	0.2235
Stdv	0.0445	0.0521	0.0578
Lolo			
Mean	0.1235	0.1056	0.1594
Median	0.1176	0.098	0.149
Stdv	0.0443	0.0462	0.0564
Bomaan			
Mean	0.1475	0.1325	0.2813
Median	0.1451	0.1255	0.2824
Stdv	0.0467	0.0551	0.0652
Berhi			
Mean	0.1178	0.1246	0.2838
Median	0.1137	0.1176	0.2824
Stdv	0.0393	0.0539	0.0477



(a)



(b)



(c)

Figure 2: A comparison between the different date varieties according to each color (RGB) (a, b, c).

As shown in figure 2, the comparison between color ingredients gave little hope to use neither red nor green to differentiate between the investigated date varieties. Blue frequency distribution demonstrated more disseminated allocation along the color intensity axis. That means the blue color is the most promising color ingredient to be used as a segregation criteria.

As a result of the comparison between red components of the five date varieties, an ANOVA test was run to find out if there are significant differences between every two-date varieties as well as the differences within each variety.

Table 2 displays the results of the ANOVA test. Shown color (R, G or B) represents the color where there is significant difference between

the two date varieties located in the corresponding vertical and horizontal cells.

As expected, the blue components can be used to separate between date varieties under investigation. There were no significant differences between Khalas and Bomaan regarding any color component. On the other hand, all color ingredients were significantly different between Fard and Bomaan .

The relationship between Sugar content and color intensity

To answer the question whether sugar content is related to color component intensity or not, Sucrose, Glucose, and Fructose content of each sample were determined using the chromatography. Each date variety was studied separately to calculate the correlation factor between each sugar content (Sucrose, Fructose, and Glucose) and the intensity of each color ingredient mean (Red, Green and Blue).

Table 2 ANOVA results between the investigated date varieties

Fard	Lolo	Khalas	Bomaan	Berhi	
B,G	B	B	B		**Berhi**
R,G,B	B	N.S.			**Bomaan**
B	B				**Khalas**

Lolo Sugar Content

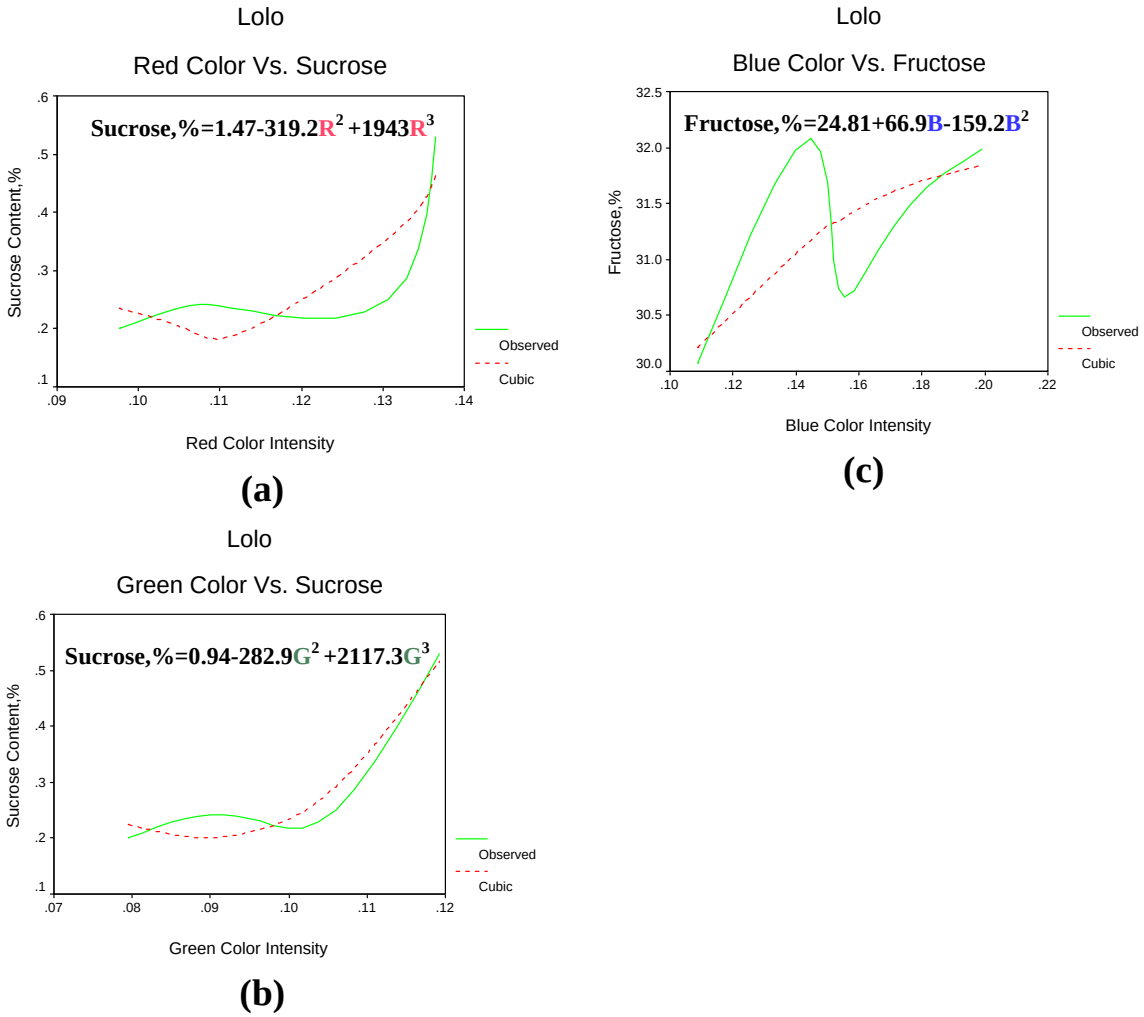
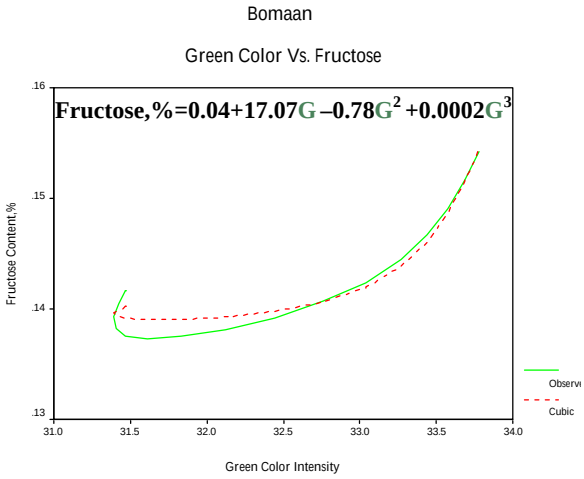


Figure 3: the relationship between Lolo color ingredients and sugars contents (a, b, c).

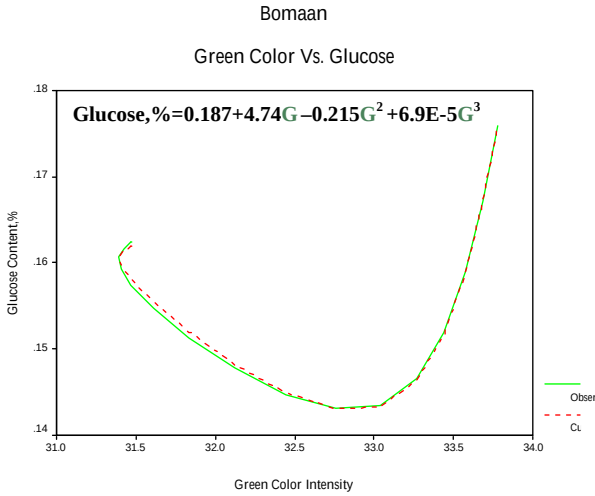
According to correlation analysis, a prediction equation was developed for the relation between red and green intensities and sucrose content. Also, the prediction equation of fructose estimation as a factor of blue color intensity was developed.

Bomaan Sugar content

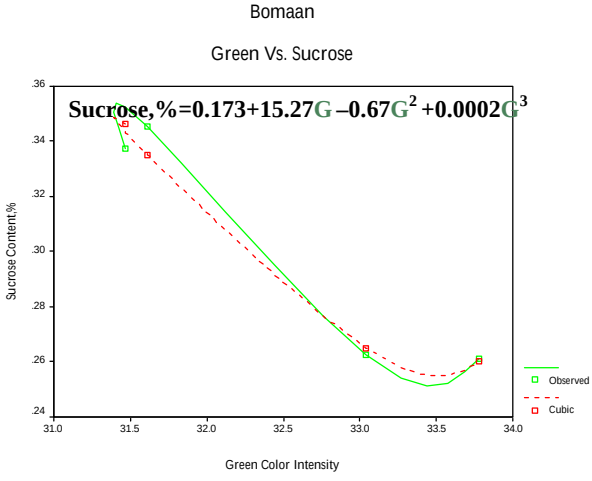
(a)



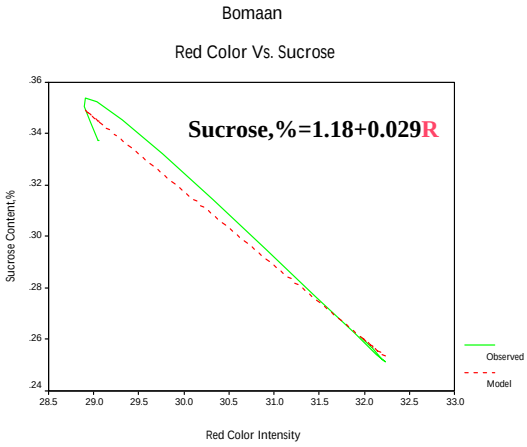
(b)

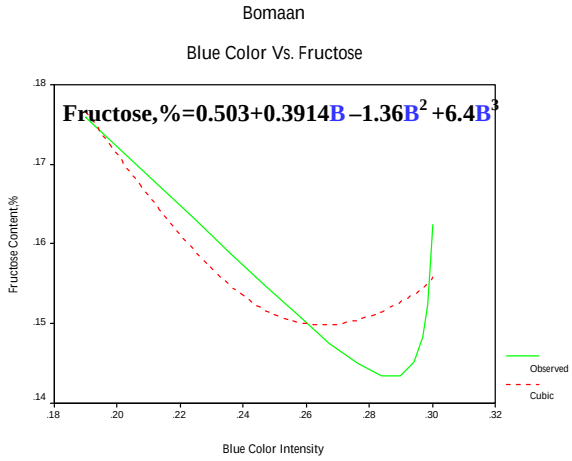


(c)

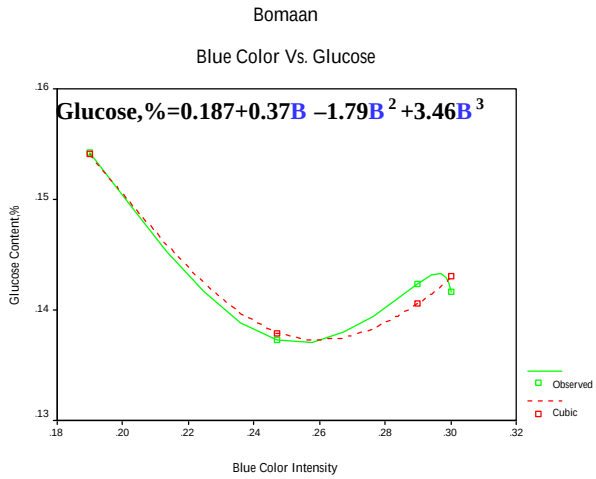


(d)





(e)



(f)

Figure 4: the relationship between Bomaan color ingredients and sugars contents (a, b, c, d, e, f).

As shown in figure 4 , fructose content may be estimated as a function of green and blue intensities. Also, glucose content could be expected according to the same colors intensities. On the other hand, green and red intensities are more promising in sucrose estimation.

Berhi sugar content

Glucose is correlated significantly to the green intensity in a way the resulted data were used to develop a prediction equation to estimate glucose content as a function of green intensity.

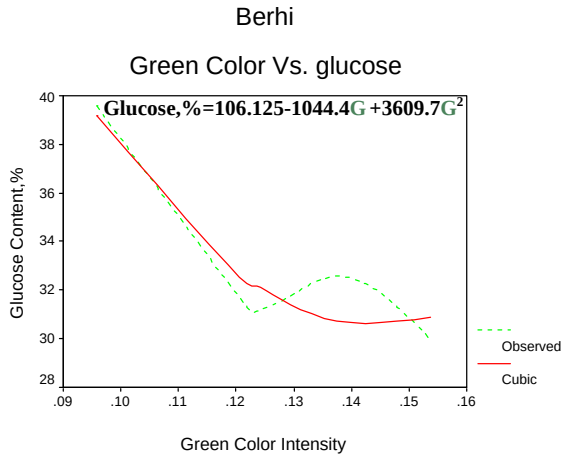


Figure 5: the relationship between Berhi green ingredient and glucose content.

Khalas sugar content

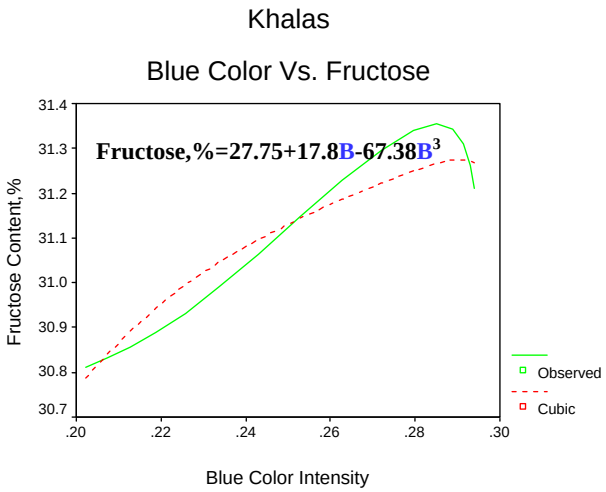


Figure 4: the relationship between Khalas Blue ingredient and fructose content.

Fard sugar content

The correlation between color intensity and sugar content did not prove its significance for the studied samples.

Conclusion

For human eyes, color differences may not be the best criteria to differentiate between date varieties. Digital imaging equipment are used to respect images as a mathematical format. Blue ingredient can be used as separation factor among investigated date varieties. Except between Bomaan and Khalas where more complicated mathematics may be needed. Also, color components (R, G, B) may be used to segregate between Fard and Bomaan. To differentiate between Fard and Berhi, Blue and green ingredients are the best factors to use. As a result, the correlation factor between sugar contents and color intensity varied according to date variety. For Lolo, it was 0.84 between sucrose content and green intensity where it was –0.017 between glucose content and blue intensity while it was –0.99 between glucose content and blue intensity of Bomaan variety and was 0.04 between fructose content and red intensity in the same variety. Also, glucose content in Berhi correlated to its green intensity with factor of –0.46 while it was –0.2 between sucrose content and red intensity. A group of prediction equations were developed in order to estimate sugar content of each variety according to its color properties.

REFERENCES

Al-Hooti, S. and Sidhu, J.S. (1997). Objective Color Measurement of Fresh Date Fruits and Processed Date Products. J. of Food Quality, Food & Nutrition Press. July 1997.v.20(3)p.257-266.

AL-Janobi, A.A. (1998) . Color Line Scan System For Grading Date Fruits. ASAE, Annual Mtg. Orlando, Florida, USA, 12-16 July, ASAE Paper # 983028.

Al-Janobi, A.A. (2000). Date Inspection by Color Machine Vision. J. King Saud Univ., v.12,Agric. Sci.(1).p.69-79.

Adamsen, F.J.; Pinter, P.J.; Barens, E.M.; LaMorte, R.L.; Wall, G.W.; Leavitt, S.W.; and Kimball, B.A. (1999) Measuring Wheat Senescence With a Digital Camera. Crop Sc. Am. May/June 1999. v.39(3) p.719-724.

Davies, J. and Perkins, R.M. (1991). Effect of Illumination in Grading Dates. ASAE, No. 91-3547.

Williams, E.J. and Adams, S.D.(1996) . Machine Vision System For Inspection of Agricultural Commodities. USDA Patents. Wash. D.C. Jan. 30, 1996 .

Wulfsohn, D.; Sarig, Y. and Algazi, R.V.(1989). Preliminary Investigation to Identify Parameters for Sorting of Dates by Image Processing. ASAE paper #89-6610