Journal of Agricultural and Marine Sciences Vol. 22 (1): 36-41
DOI: 10.24200/jams.vol22iss1pp36-41
Reveived 10 Jun 2016
Accepted 15 Nov 2016

RESEARCH PAPER

Computer vision technique to classify dates based on hardness

Naeema H. Al-Shekaili[1], Annamalai Manickavasagan[2*], Nawal K. Al-Mezeini[2],
M. Shafiur. Rahman[1] and Negib. Guizani[1]

استخدام تقنية تصوير الحاسوب في تصنيف التمور حسب الصلابة

نعيمة حارب الشكيلي[1] ومنيكافا سجان أنامالي[2*] ونوال خميس المزيني[2] ومحمد شفير رحمان[1] ونجيب قيزاني[1]

Abstract. Hardness is one of the important attributes in determining the quality of dried fruits. Hardness assessment is normally carried out by manual inspection. This method is time consuming, laborious, expensive and subjective. The objective of this study was to develop a computer vision system with a monochrome camera to classify dates based on hardness. Date samples were obtained from three different growing regions in Oman and graded into soft, semi-hard, and hard classes based on hardness. A total of 1800 date samples were imaged individually using a monochrome camera (600 dates / class). Histogram and texture features were extracted from the acquired monochrome images and used in the classification models. The overall classification accuracies in three class model (soft, semi-hard, and hard) were 66% and 71% for linear discriminant analysis (LDA) and artificial neural network (ANN), respectively. It was improved to 84% and 77% in LDA and ANN, respectively while using two class model (soft and hard (semi-hard and hard together)). The histogram features were more contributing in the date classification based on hardness than image texture features. Computer vision technique has great potential to develop online quality monitoring systems for dates and other dried fruits.

Keywords: Dates; hardness; histogram features; texture features; gray scale images

الملخص: تعتبر الصلابة من أهم الخصائص في تقييم جودة الفواكة الجافة. ويتم تقييم الصلابة عادة عن طريق التفتيش اليدوي، إلا أن هذه الطريقة تتطلب الجهد والوقت الطويل، كما أنها باهضة وغير موضوعية. تهدف هذه الدراسة إلى تطوير أسلوب التصوير باستعمال الحاسوب متصلا بكاميرا تصوير أحادية اللون بهدف تصنيف التمور اعتماداً على الصلابة. وقد تم الحصول على عينات التمور من ٣ مناطق مختلفة بسلطنة عمان وصنفت هذه التمور إلى ٣ فئات: لينة، شبه صلبه، وصلبه. حيث تم تصوير مجموع ١٨٠٠ عينة (٦٠٠ عينة لكل فئة) باستعمال كاميرا التصوير أحادية اللون. وقد تم استخراج ملامح قوام التمور من الرسم البياني واستخدمت كنموذج تصنيف. وقد حققت إجمال تصنيف التمور باستعمال نموذج الثلاث فئات (اللينة، شبه صلبه، وصلبه) دقة تقدر بـ٦٦٪ و ٧٩٪ عند استعمال طريقة التحليل الخطي التمييزي (LDA) و طريقة التحليل الخطي المتدرج (ANN) على التوالي. كما حققت النتائج نسب دقة أعلى عند استعمال نموذج الفئتين (فئة لينة، وفئة شبه صلبه، وصلبه معا) تقدر بـ ٨٤٪ و ٧٧٪ بطريقة (LDA) و(ANN) على التوالي. وقد كانت مساهمة ملامح الرسم البياني أوضح من ملامح القوام في تصنيف التمور. يمتلك التصوير باستعمال الحاسوب قدرات عالية يمكن استخدامها لتطوير أنظمة مراقبة جودة التمور والفواكة الجافة الأخرى عبر الأنترنت.

الكلمات المفتاحية: التمور ، الصلابة ، ملامح الرسم البياني ، ملامح القوام ، صور المقياس الرمادي

Introduction

Date palm (*Phoenix dactylifera*, L.) is one of the oldest and most important staple crops in the Middle East and North Africa. The Sultanate of Oman is ranked among the top ten date producing countries in the world with approximately 49% of cultivable land area (FAO 2010). Although date production was 276,400 MT in 2010 (FAO, 2010), only about 5,000 tons was exported (Zaabanot, 2011). The low level of export from Oman is due to improper sorting of dates to ensure higher quality as expected by consumers (Al-Marshudi, 2002). According to the CODEX standard, the quality attributes to grade dates are colour, flavor (sugar level), moisture content and absence of defects, such as insect damage, and surface damage (Kader and Hussein, 2009).

Moisture content in dates is important because excessive loss of water causes drying, consequently they becomes hard (Kader and Hussein, 2009; Rahman and Al-Farsi, 2005). Hardness beyond a critical value is considered as a defect in dates as it affects the physical properties and consumer acceptability. Dates can be subdivided into: soft, semi-dry (semi-hard) and dry (hard) according to their moisture content or hardness (Kader and Hussein, 2009; Al-Janobi, 1998). Hard dates are chewy and tough with strong curvy and zigzag textured skin (Rahman and Al-Farsi, 2005). Dates are processed into different products and the choice of the date types for a given product depends on the final product. For example, soft dates are used to manufacture date syr-

[2*]Annamalai Manickavasagan (✉) Sultan Qaboos University, College of Agricultural and Marine Sciences, Department of Soils, Water and Agricultural Engineering. Box 34, Al-Khod 123. Sultanate of Oman. current email: mannamal@uoguelph.ca.
[1] Sultan Qaboos University, College of Agricultural and Marine Sciences, Department of Food Science and Nutrition.

up while hard dates are used to produce flour. Therefore hardness is one of the important parameters used to evaluate and classify dates in industry. The presence of hard dates in other grades affects the acceptability of the whole batch and yields low values in domestic and international markets.

In general, hardness assessment is carried out by the traditional method of visual inspection or mechanical methods like diagonal metal plate method or vacuum systems (Huxsoll and Reznik, 1969; Chesson et al., 1979; Al-Janobi, 2000). The visual inspection method is subjective, laborious and expensive, whereas mechanical methods are sample-destructive nature and conducted only on representative samples. An objective non-destructive method for sorting of dates based on hardness would be highly beneficial for online quality assessment and monitoring of dates in handling facilities. Therefore the objective of this study was to determine the ability of a computer vision system with a monochrome camera to classify dates based on hardness. The monochrome camera was selected because of its low cost, image size and faster image handling and processing capabilities.

Computer based image processing techniques or computer vision technologies replace the traditional method of human inspection towards achieving better, faster, and automated operations (Patel et al., 2012; Pour-Damanab et al., 2012). It gives a meaningful description for the object by duplicating human vision using different algorithms to assess the quality (Narendra and Hareesh, 2010). It is used to characterize colour, texture and complex geometric properties (Chandraratne et al., 2003). Computer vision is used in food industry for the classification, quality evaluation, sorting, grading, and defect detection (Du and Sun, 2006; Brosnan and Sun, 2004). The application of computer vision in the date industry is scarce (Lee et al., 2008a). Schmilovitch et al. (1999) developed a semi-automated vision system for maturity determination of fresh dates using NIR scanner. Al-Janobi (2000) graded Saudi dates (Sifri variety) based on colour and texture with an average error of 1.8% using a colour camera. A computer vision system for the grading of dates based on fruit size and skin delamination using reflected NIR imaging showed 10% higher accuracy over human inspection and a reduction in labor cost by 75% (Lee et al., 2008b). In another study, using a RGB colour imaging system to grade dates into 3 categories based on size, shape, flabbiness intensity and defects yielded 80% accuracy (Al-Ohali, 2011).

Materials and methods

Sample collection

Samples of dates of the Fard variety, the most processed variety in Oman, were obtained from three major dates growing regions: Al-Batinah, Al-Dakhliah and Al-Sharqiah. The dates were sorted into three classes based on hardness (hard, semi-hard and soft) by an experienced grader and confirmed by a manager in Bright Sun Dates Company, Oman. From each region, a representative sample of 600 dates (200 per class), in total 1800 samples, were selected and used in this study.

Image acquisition system

A monochrome camera (model: XCD-X700, Sony, Japan) was used to acquire uncompressed 8-bit images of resolution 1024×768 pixels with a charge coupled device (CCD) sensor. It was placed at a height of 1 m in order to simulate the application in handling facilities. Images were taken in a dark room while illuminating the sample with two fluorescent lights (36 W, model: Dulux L, OSRAM, Italy). The camera was connected to the computer through IEEE 1394 cable. The camera was calibrated using white and black standard colour cards (Digital Kard XL, DGK colour Tools, USA) before starting each batch of imaging.

Image analysis and classification model

The region occupied by date sample in the monochrome image was segmented from the background using Matlab (version 7.6.0) (Fig. 1). Histogram and texture features of the grayscale information for the region pertaining to date sample were extracted and analyzed statistically (at $\alpha \leq 0.05$). The details of the extracted features are given in Table 1.

The classification accuracy of a linear discriminant analysis (LDA) and stepwise linear discriminant analysis (SLDA) was performed with SPSS software. Similarly the accuracy of a back propagation neural network (BPNN) was obtained using Matlab.

Discussion

Although date samples used in this study belonged to the same variety, regional differences existed in their

Figure 1. Process involved in the segmentation of dates from the background from left to right (a) original monochrome image (b) adjusted monochrome image (c) binary image (d) binary image after the morphological operations to reduce noise (e) segmented date.

Table 1. Features extracted from monochrome images.

Features	Definition*
Histogram features	
Mean gray value	Average of the gray values of all the pixels in an image
Standard deviation	Standard deviation of all the pixels in an image
Variance	Variance of all the pixels in an image
Smoothness	Measure of the relative smoothness of the intensity in a region
Eccentricity	Ratio of distance between the foci of the ellipse and its major axis length
Solidity	Proportion of the pixels in the convex hull that are in a region
Extent	Proportion of the pixels in the bounding box that is in the region
Texture features (GLCM)**	
Contrast	Measure of contrast between a pixel and its neighbor over the whole image
Correlation	Measure of how correlated a pixel is to its neighbor over the whole image
Energy	Sum of squared elements in the GLCM
Homogeneity	Similarities of pixels
Maximum probability	Maximum occurrence of the gray level
Entropy	Measure of the randomness of intensity image
Cluster prominence	Measure of the skewness of a matrix
Cluster shade	Measure of the lake of symmetry
Dissimilarity	Measure of the dissimilarity between the pixels

* Manickavasagan et al. (2008a,b); Basavaraj and Vishwanath (2009); Gonzalez et al. (2010)
** Gray Level Co-occurrence Matrix

external properties such as colour, shape and size. In general, hard dates were brighter in colour compared to soft and semi-hard dates. However, there were several overlaps in gray scale values across different classes and regions and interfered in the classification.

Features extracted from monochrome images

Histogram features

The histogram features of the date samples from different regions are shown in Table 2. There was a difference between soft, semi-hard and hard dates in the mean gray value. The maximum and minimum mean gray values were associated with hard and soft dates, respectively. This indicated that the hard dates were brighter in colour compared to the soft and semi hard dates. There was no difference in the mean gray value of the dates from the Al-Dakhliah and Al-Batinah regions. Standard deviation and variance of the soft, semi-hard and hard dates varied significantly. It was the highest for soft dates and the lowest for hard dates. There was no difference in the standard deviation and variance of the dates from Al-Batinah and Al-Sharqiah regions.

The smoothness was different between the three classes of dates with maximum and minimum values for soft and hard dates, respectively. However, there was no difference in smoothness between dates from the Al-Sharqiah and Al-Dakhliah regions. There was a difference in eccentricity between three classes with hard dates having the highest values and soft dates the lowest values. The eccentricity of the dates was not different between the Al-Batinah and Al-Sharqiah regions. The growing regions produced differences in solidity of date samples. The extent of three classes of dates was significantly different. The highest and lowest value was obtained for hard and soft dates, respectively.

Table 2. Mean values of histogram features extracted from monochrome images of dates (n=200).

Feature	Region								
	Al-Batinah			Al-Dakhliah			Al-Sharqiah		
	Soft	Semi-hard	Hard	Soft	Semi-hard	Hard	Soft	Semi-hard	Hard
Mean gray value	45.69	48.34	57.16	45.88	49.03	56.38	47.50	53.84	61.91
Standard deviation	39.26	37.34	33.44	36.07	35.08	33.98	39.62	36.22	32.90
Smoothness	0.9993	0.9992	0.9990	0.9992	0.9991	0.9990	0.9993	0.9992	0.9989
Eccentricity	0.7754	0.7820	0.7958	0.7950	0.8127	0.8089	0.7633	0.7861	0.7959
Solidity	0.9867	0.9871	0.9866	0.9862	0.9861	0.9864	0.9848	0.9861	0.9861
Extent	0.7996	0.8047	0.8124	0.8096	0.8134	0.8164	0.7959	0.8048	0.8160

Table 3. Mean values of texture features extracted from monochrome images (n=200).

Feature	Region								
	Al-Batinah			Al-Dakhliah			Al-Sharqiah		
	Soft	Semi-hard	Hard	Soft	Semi-hard	Hard	Soft	Semi-hard	Hard
Contrast	0.0246	0.0241	0.0232	0.0235	0.0230	0.0240	0.0222	0.0221	0.0215
Correlation	0.9956	0.9956	0.9956	0.9956	0.9957	0.9958	0.9951	0.9953	0.9953
Energy	0.8839	0.8859	0.8893	0.8876	0.8876	0.8789	0.9060	0.9017	0.9038
Homogeneity	0.9995	0.9995	0.9995	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996
Maximum probability	0.9381	0.9392	0.9411	0.9402	0.9402	0.9352	0.9505	0.9482	0.9493
Entropy	0.2349	0.2317	0.2262	0.2291	0.2291	0.2424	0.1992	0.2063	0.2027
Cluster prominence	1824.32	1798.46	1754.3	1779.45	1779.99	1883.85	1532.95	1592.93	1562.43
Cluster shade	138.34	136.25	132.71	134.66	134.70	143.259	114.96	119.70	117.34
Dissimilarity	0.00351	0.00344	0.00331	0.003367	0.0032	0.00343	0.00318	0.00316	0.00307

Texture features

The values of texture features for dates are given in Table 3. The soft dates had more contrast than the semi-hard and hard dates. However, there was no difference in the contrast of semi-hard and hard dates. It was maximum for Al-Batinah and minimum for Al-Sharqiah regions. The correlation of dates varied significantly with respect to classes and regions. There was no difference in the energy between the three classes. However, there was a difference in the energy between the three regions. The homogeneity of date samples from three regions was different with maximum and minimum values for Al-Sharqiah and Al-Batinah regions, respectively. However, the homogeneity of hard and semi-hard dates was not different. There was no difference in the maximum probability between three classes of dates. On the other hand, the maximum probability between the three regions was significantly different. The Al-Sharqiah region had the highest and Al-Dakhliah had the lowest probability. The entropy, cluster prominence and cluster shade of the date samples from three regions were different. The Al-Dakhliah region had the highest entropy, cluster prominence and cluster shaded while Al-Sharqiah had the lowest values. However, there was no difference in the entropy, cluster prominence and cluster shaded between the three classes of dates. The soft dates were more dissimilar in comparison to semi hard and hard dates. But there was no difference in the dissimilarity of semi-hard and hard dates.

Table 4. Classification accuracy (%) of date samples in three and two class models.

Region	LDA*	SLDA**	Selected features for SLDA
Three-class model			
Al-Batinah	59	58	mean gray value, variance and extent
Al-Dakhliah	67	66	mean gray value, eccentricity, smoothness, cluster prominence and maximum probability
Al-Sharqiah	76	75	mean gray value, variance, extent, correlation and smoothness
All regions together	66	66	mean gray value, standard deviation, variance, extent, correlation, smoothness, dissimilarity and maximum probability
Two-class model			
Al-Batinah	83	83	mean gray value, variance and extent
Al-Dakhliah	87	87	mean gray value, variance, solidity, extent , smoothness, cluster prominence and maximum probability
Al-Sharqiah	86	85	mean gray value, standard deviation and extent
All regions together	83	84	mean gray value, standard deviation, solidity, extent, smoothness, entropy, maximum probability

* all sixteen features were used in the linear discriminant analysis (LDA)

** most contributing features were used in the stepwise linear discriminant analysis (SLDA)

Table 5. Accuracies (%) of ANN to classify dates according to hardness in the 3-class model.

From / To	Soft	Semi-hard	Hard
Soft	76	22	2
Semi-hard	17	68	15
Hard	5	25	70

Table 6. Accuracies (%) of ANN to classify dates according to hardness in the 2-class model.

From / To	Soft	Hard
Soft	69	31
Hard	16	84

Classification models

Linear discriminant analysis (LDA)

The features extracted from the monochrome images were used to determine the efficiency of this technique in sorting of dates based on hardness. In the first approach, the date samples were classified into three groups namely soft, semi-hard and hard (three class model). In some applications, dates are graded into only two categories such as soft and hard. Therefore in the second approach, the date samples were classified into two groups namely soft and hard (semi hard and hard together) (two class model). Analyses were also performed by considering each region separately and combined together.

Table 4 shows the accuracies obtained in different approaches. In three class models, it was in the range of 58% to 76%. The highest and lowest classification accuracies were obtained for the Al-Sharqiah and the Al-Batinah regions, respectively. While analyzing all regions together, there was no difference between LDA and stepwise linear discriminant analysis (SLDA) with most contributing factors. The SLDA selected mostly histogram features for the classification. This indicates that histogram features are more contributing than the texture features in the classification of dates based on hardness. Similarly, Basavaraj and Vishwanath (2009) reported that texture features including contrast, correlation, energy, entropy, homogeneity and dissimilarity were not sufficient for the classification of bulk sugary foods. On the other hand, Chandraratne (2003) reported that image texture features were suitable indicators for beef tenderness because the R2 value was 0.621 while using geometrical features and 0.746 with texture features. Also, Li et al. (1999) and Li et al. (2001) mentioned the same about the importance of texture features in classification of beef tenderness.

In two class models, the classification accuracy was 83% to 87%. The highest classification in this approach was achieved for Al-Sharqiah region.

Zayas et al. (1996) obtained 63% accuracy for hard wheat and 91% accuracy for soft wheat in two class model, using monochrome images. Li et al. (2001) classified the steaks using colour camera into rough and tender with an accuracy of 83%.

Artificial neural network (ANN)

The classification accuracies of ANN for three class model are shown in Table 5. The misclassification was observed between the soft and semi-hard and hard and semi hard dates. The overall accuracy obtained was 71%. Fadel (2007) used colour camera and obtained a classification accuracy of 100%, 80%, 80%, 60% and 80% for Fard, Khalas, Lolo, Bomaan and Berhi dates varieties, respectively using probabilistic neural network. While grading the dates into three grades according to size, shape, flabbiness intensity, and defects using RGB images, Al-Ohali (2011) obtained 55% to 80% classification accuracy.

In two class model, the classification accuracy of the hard dates was higher than soft dates (Table 6). The overall accuracy in the two class model was 77%.

Conclusion

A computer vision system with monochrome camera was developed to classify dates based on hardness with varying degree of success. The classification was carried out with histogram and texture features extracted from the monochrome images of dates using LDA and ANN. The classification accuracies of two class models were higher than three class models in both LDA and ANN. In SLDA, histogram features contributed more for classification than texture features. The potential of computer vision technique for hardness determination in dates must be investigated with other cameras such as NIR and RGB colour cameras to improve classification accuracy.

Acknowledgements

The authors would like to acknowledge The Research Council (TRC) in Sultanate of Oman for funding this work under the project *"Development of computer vision technology (CV) for the quality assessment of dates in Oman"*.

References

Al-Janobi, A. 1998. Application of co-occurrence matrix method in grading date fruits. Paper No. 98 – 3024. ASAE meeting presentation, King Saud University.

Al-Janobi, A. 2000. Date inspection by color machine vision. Journal of King Saud University of Agricultural Science 12: 69 – 79.

Al-Marshudi, A.S. 2002. Oman traditional date palms: Production and improvement of date palms in Oman. Tropicultura 20: 203 – 209.

Al-Ohali, Y. 2011. Computer vision based date fruit grading system: Design and implementation. Journal of King Saud University, Computer and Information Sciences 23: 29 - 36.

Basavaraj, S.A. and Vishwanath, C. B. 2009. Texture based identification and classification of bulk sugary food objects. ICGST-GVIP J 9: 9-14.

Brosnan, T. and Sun, D. 2004. Improving quality inspection of food products by computer vision-a review. Journal of Food Engineering 61: 3-16.

Chandraratne, M.R., Samarasinghe, S., Kulasiri, D., Frampton, C., Bekhit, A.E.D. and Bickerstaffe, R. 2003. Prediction of lamb tenderness using texture features. Research report, Lincoln University, ISSN 1174 - 6696.

Chesson, J.H., Burkner, P.F. and Perkins, R.M. 1979. An experimental vacuum separator for dates. Transactions of ASAE, 22: 16 - 20.

Du, C. and Sun, D. 2006. Learning techniques used in computer vision for food quality evaluation: a review. Journal of Food Engineering 72: 93 - 55.

Fadel, M. 2007. Date fruits classification using probabilities neural networks. Agricultural Engineering International: The CIGR Journal (CIGR Commission Internationale du Genie Rural J). Manuscript ID 07 003. IX: 1-10.

FAO. 2010. http://faostat.fao.org/site/339/default.aspx. Access done on 14th January 2013.

Gonzalez, R.C., Wood, R.E. and Eddins, S.L. 2011. Digital image processing using MATLAB. New Delhi: Tata McGraw Hill Education Private Limited.

Huxsoll, C.C. and Reznik, D. 1969. Sorting and processing mechanically harvested dates. Date Grower's Institute Report 46: 8 - 10.

Kader, A. and Hussein, M. 2009. Harvest and postharvest handling of dates. Project on the development of sustainable dates palm production system in the GCC countries of Arabian Peninsula. International Center for Agricultural Research in the Dry Areas.

Lee, D., Archibald, J. , Chang, Y. and Gerco, R. 2008a. Robust color space conversion and color distribution analysis technique for date maturity evaluation. Journal of Food Engineering 88: 364 - 372.

Lee, D., Schoenberger, R., Archibald, J. and McCollum, S. 2008b. Development of a machine system for automatic date grading using digital reflective near-infrared imaging. Journal of Food Engineering: 388 - 398.

Li, J., Tan,J., Martz, F.A. and Heymann, H. 1999. Image texture feature as indicators of beef tenderness. Meat Science 53, 17 - 22.

Li, J., J. Tan and P. Shatadal. 2001. Classification of tough and tender beef by image texture analysis. Meat Science 57: 341 - 346.

Manickavasagan, A., Sathya, G. and Jayas,D.S. 2008a. Comparison of illumination to identify wheat classes using monochrome images. Computer and Electronics in Agriculture 63: 237-244.

Manickavasagan, A., Sathya, G., Jayas, D.S. and White, N.D.G. 2008b. Wheat class identification using monochrome images. Journal of Cereal Science 47: 518-527.

Narendra, V.G. and Hareesh, K.S. 2010. Quality inspection and grading of agricultural and food products by computer vision – a review. International Journal of Computer Application 2: 43-65.

Patel, K.K., Kar, A., Jha, S.N. and Khan, M.A. 2012. Machine vision system: a tool for quality inspection of food and agricultural products. Journal of Food Science and Technology 49: 123-141.

Pour-Damanab, A.S., Jafary, A. and Rafiee, S. 2012. Kinetics of the crust thickness development of bread during baking. Journal of Food Science and Technology, doi: 10.1007/s13197-012-0872-z.

Rahman, M.S. and Al-Farsi, S. 2005. Instrumental texture profile analysis (TPA) of date flesh as function of moisture content. Journal of Food Engineering 66: 505 - 511.

Schmilovitch, Z., Hoffman, A., Egozi, H., Ben-Zvi, R., Bernstin ,Z. and Alchanatis,V. 1999. Maturity determination of fresh dates by near infrared spectrometry. Journal of the Science of Food and Agriculture 79: 86 - 90.

Zaabanot, A. 2011. Improve date syrup production in Oman – Case study. Published in Oman Daily, Saturday 8th October, Edition No11085: 22.

Zayas, I.Y., Martin, C.R., Steele, J.L. and Katsevich, A. 1996. Wheat classification using image analysis and crush-force parameters. Transactions of the ASAE, 39: 2199 - 2204.