Journal of Food Engineering 128 (2014) 72–78

Contents lists available at ScienceDirect

Journal of Food Engineering

journal homepage: www.elsevier.com/locate/jfoodeng

A digital image analysis method for assessment of lentil size traits

L.S. LeMasurier *, J.F. Panozzo, C.K. Walker

Department of Environment and Primary Industries, PB260 Horsham, Victoria 3401, Australia

ARTICLE INFO

Article history:
Received 23 July 2013
Received in revised form 1 November 2013
Accepted 19 December 2013
Available online 28 December 2013

Keywords:
Digital Image Analysis (DIA)
Machine vision
Lentil
Seed Size Distribution (SSD)
Seed Size Index (SSI)
Quality assessment

ABSTRACT

Image analysis offers a rapid and non-destructive method for a wide range of agricultural applications including grading of grains. In this study, a method was developed for seed size grading by analysis of three-dimensional digital image information on single lentil kernels. Predicted length, single-seed mass, bulk-sample mass and Seed Size Index (SSI) were all highly correlated with the physically measured values giving Lin's concordance test statistics of 0.98, 0.97, >0.99 and >0.99 respectively. Sieves were found to have a 10% misclassification rate and Seed Size Distribution was predicted within error of the sieving method. Results also indicated that image analysis can give much more detailed and precise descriptions of grain size and shape characteristics than can be practically achieved by manual quality assessment

1. Introduction

Pulse grains are an important component of cropping systems. Sown in rotation with cereal and oilseed crops, pulses act as a disease break and increase the productivity by improving soil nutrient levels (Nleya et al., 2004). Moreover, by increasing soil fertility, pulse grains lessen the need for chemical and fertiliser inputs. However, farmers can be deterred from growing pulses as quality is largely based on appearance of the grain and assessed through subjective methods and as a result farmers may not achieve the true market price.

Seed size is a major indicator of quality in lentils. It is desirable to have minimal variation in seed size within a cultivar as uniformity enhances efficiency of post-farm grain processing, particularly de-hulling and splitting. Currently the standard for determining Seed Size Distribution (SSD) of a pulse grain sample is to sieve the sample through a series of nested, round-hole screens with aperture varying in diameter from 3 mm to 7 mm, and calculate percentage mass of grain in each size fraction. Seed Size Index (SSI), the expected value of SSD, is then taken as the representative grain size for a sample. However, this mechanical sieving process is tedious, time consuming and error prone (Shahin and Symons, 2005).

Digital Image Analysis (DIA) provides a rapid, objective and non-destructive method, for replicating manual and visual quality-assessment techniques, which is relatively inexpensive (Brosnan and Sun, 2002). Image based grading of foodstuff has been investigated for a wide range of applications over the past three decades (Brosnan and Sun, 2004; Costa et al., 2011; Tillett, 1991) and is now commonplace across the agricultural industry (Kondo, 2010). Within the grains sector, methods for quality assessment and grain classification by DIA are continually being developed and applied particularly for cereal grains (Davies, 2009; Visen et al., 2004; Walker and Panozzo, 2012) and increasingly for other grain types such as pulse grains (Fıratlıgil-Durmuş et al., 2010; Shahin and Symons, 2001) and rice (Aggarwal and Mohan, 2010; Yoshioka et al., 2007).

Ideally digital-imaging quality-assessment systems would be applicable in-field and across industry as well as in laboratory environments. Under field and industry conditions, many proposed imaging systems, particularly those with three-dimensional capabilities, are currently impractical and so two-dimensional imaging technologies have been widely utilised for size predictions of grains and granular material (Mandal et al., 2012; Murtagh et al., 2005; van Dalen, 2004; Walker and Panozzo, 2012). Results of these studies support the efficacy of DIA as a replacement for manual size grading, however by collecting only two-dimensional information the third dimension must be assumed. Typically, scanners have been used for single seed analysis, which requires careful placement of grain so that all kernels are unconnected. This can be time consuming and restrictive on sample size. Shahin and Symons (2005) developed DIA algorithms for performing seed sizing of non-singulated kernels on a scanning device, however in this approach some information on size and shape is lost by overlapping grains. Imaging systems comprising a camera unit positioned above a conveyor belt, have been successfully employed for analysis of fruits and vegetables much

* Corresponding author. Tel.: +61 3 5362 2170; fax: +61 3 5362 2198.
E-mail address: linda.lemasurier@depi.vic.gov.au (L.S. LeMasurier).

0260-8774/$ - see front matter
http://dx.doi.org/10.1016/j.jfoodeng.2013.12.018

larger in size than grains (Chen and Ting, 2004; Cubero et al., 2011; ElMasry et al., 2012).

Lentils are among the simplest of all pulse grains, in characterisation of size and shape through DIA, due to their lesser geometric complexity by comparison. Therefore, image-based methods for sizing lentil samples provide a good basis to extend into a much more versatile pulse quality-assessment tool. The recent development of imaging systems such as the EyeFoss™ (FOSS Analytical, Höganäs, Sweden), has provided the hardware for practical, efficient, and precise quality analysis of grains in industry as well as laboratory environments. In this study an EyeFoss™ was used to capture three-dimensional single seed information, Red–Green–Blue (RGB) digital images and laser measurements of height, for constructing lentil size grading algorithms. This research describes the development of image analysis algorithms for efficient and accurate predictions of lentil size quality traits; length, mass, SSD and SSI and the applications of these algorithms to inform plant breeding programs.

2. Materials and methods

2.1. Samples and manual seed sizing

Lentil samples were taken from several red, green and red-green crossed lentil varieties from the 2009/2010, 2010/2011 and 2011/2012 growing seasons. These were provided by the Victorian Department of Environment and Primary Industries, Pulse Breeding and Agronomy groups. Seed sizing was done manually on clean samples (i.e. no pods or non-seed material) using industry-standard, nested, round-hole sieves (7.0, 6.0, 5.0, 4.5, 4.0, 3.5 and 3.0 mm) and all material smaller than 3.0 mm was discarded as this is not included in standard calculations of SSI. Lentils sized as 3.0 mm or larger were used to find the SSD and SSI of each sample by industry standard methods. Seeds retained through the manual sieving process were also imaged. Sample size ranged between approximately 1000 and 3500 seeds for bulk samples.

2.2. Image acquisition and pre-processing

An EyeFoss™ (FOSS Analytical, Höganäs, Sweden), comprising a fixed-resolution, line-scan camera and laser (Ranger E, SICK Sensor Intelligence) mounted over a conveyor system, with Foss Integrator Software version 1.5.3, was used to simultaneously capture digital images and measure height of lentils. Sample presentation involved loading seed into a hopper which was mechanically lifted dispensing seeds onto a vibrating comb and onto the conveyor belt so that the seeds were disjoint. The seeds were then carried by conveyor under the camera unit and imaged individually. Occasionally seeds would overlap on the conveyor belt and so some images would contain multiple seeds. Images were automatically rotated by inbuilt EyeFoss algorithms such that the longest detected seed dimension ran vertically in the image. Images were stored as Portable Network Graphics (PNG) files and analysed in the Matlab R2012a programming environment with the Image Processing Toolbox.

Each digital image was imported into Matlab in RGB colour space and transformed also into Hue-Saturation-Value (HSV) colour space. Image segmentation, to isolate the seed in each image, was performed by setting thresholds on the H and S components of each HSV image and also on the B component of each RGB image to create a binary mask of the seed. The seed boundary, defined as the set of pixels which are within the Region of Interest (ROI) and neighbour at least one background pixel, was detected through use of the *bwtraceboundary* function in the Matlab Image Processing Toolbox. Image segmentation and boundary detection is detailed and illustrated further in Appendix A.

One kernel was analysed per image, so in the case that multiple seeds were detected in an image, the largest isolated region was chosen as the seed for size analysis. However, if the area of the largest region exceeded 12000 pixels the image was discarded without further analysis as this was considered to be too large to represent an individual lentil. The proportion of images discarded did not impact significantly on size analyses of bulk-seed samples.

2.3. Length calibration

Over 150 rectangular balsa wood blocks varying in length and width (from 2.53 to 21.72 mm) and varying in height (from 1.66 to 10.11 mm) were imaged twice in each of two orientations; first aligned parallel and then perpendicular to conveyer belt motion. Image block dimensions of length and width, measured in pixels, were found using the ruler in the interactive *imtool* function of the Image Processing Toolbox. Physical block dimensions, in mm, were measured with digital callipers. Block height, in mm, was measured automatically by laser (through the EyeFoss™) and manually by digital callipers.

Distortion of lengths was observed in the calibration block images due to both conveyor belt motion and the camera's fixed-resolution. Pixel block-lengths parallel to conveyor belt motion were compressed and lengths perpendicular were stretched. And images rotated such that the longest diameter ran vertically in the image. As the camera resolution and conveyer speed came pre-set by the manufacturer it was necessary to correct images to account for this. The extent of this stretching and compressing was found to be linearly dependent on height of the block.

$$\begin{bmatrix} x_{m1} & y_{m1} \\ x_{m2} & y_{m2} \\ \vdots & \vdots \\ x_{mn} & y_{mn} \end{bmatrix} = \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \vdots \\ x_n & y_n \end{bmatrix} RD$$

where $R = \begin{bmatrix} \cos(\theta_R) & -\sin(\theta_R) \\ \sin(\theta_R) & \cos(\theta_R) \end{bmatrix}$ and $D = \begin{bmatrix} x_{dil} & 0 \\ 0 & y_{dil} \end{bmatrix}$

A matrix transformation (Eq. (1)), to convert pixel coordinates, (x_i, y_i), to metric units, (x_{mi}, y_{mi}), was constructed in two components; rotation (R) and scaling (D). Pixel coordinates were rotated about the centre of the image, by an angle θ_R, to undo automated image rotation by the EyeFoss™. The coordinates were then scaled in directions both parallel and perpendicular to conveyor belt motion by factors x_{dil} and y_{dil} respectively. The scaling transformations (Eqs. (2) and (3)) were constructed through multiple linear regressions (MLR), where the dependent variable was length (mm) measured by callipers and the independent variables were laser-measured height (mm) and image block-length (pixels).

$$length\_parallel_{mm} = 0.0632^{*} length\_parallel_{pixels} + 0.0214^{*} height - 0.0444 \quad (2)$$

$$length\_perpendicular_{mm} = 0.0596^{*} length\_perpendicular_{pixels} - 0.0694^{*} height + 0.2790 \quad (3)$$

2.4. Length validations

A set of 100 lentil seeds of diverse sizes and shapes were chosen from several lentil varieties and used to validate DIA length predictions. Each seed was marked across a diameter, which was then measured (in mm) with digital callipers and imaged through the